PRODUCTION OF INREASED OIL AND PROTEIN IN PLANTS BY THE DISRUPTION OF THE PHENYLPROPANOID PATHWAY

This application claims priority to U.S. Provisional Application 60/427,313, filed 11/18/2002, herein incorporated by reference in its entirety.

The present invention relates to the fields of nucleic acid chemistry and agricultural biotechnology. In particular, the present invention is directed at methods and compositions for increasing oil levels in plants.

Currently, levels of oil in oilseed crops have been increased incrementally by traditional breeding and selection methods. There exist few references to transgenic plants with increased levels of oil. In contrast, increases in the proportions of some strategic fatty acids have been achieved by the introduction or manipulation of various plant fatty acid biosynthesis genes in oilseeds. For instance, Voelker *et al.* demonstrated that expression in *Brassicaceae* of a medium chain fatty acyl-ACP thioesterase from California Bay increased the lauric acid (12:0) content (*Science*, 257:72-74 (1992)). Hitz *et al.* increased proportions of oleic acid in *Glycine max* by co-suppression using a sense construct encoding a plant microsomal FAD-2 (delta-12) desaturase (Proc. 9th International Cambridge Rapeseed Congress UK, pp. 470-472 (1995)). Although the use of plant transgenes resulted in altered proportions of sn-2 lauric acid in canola and oleic acid in soy, there was no evidence of increased total fatty acid content, or increased oil yield in these transgenics.

Certain workers have attempted to increase or modulate the oil content of plants by manipulation of oil biosynthetic pathway genes. For example, U.S. Patent 6,268,550 to Gengenbach *et al.* provides maize acetyl CoA carboxylase nucleic acids for altering the oil content of plants. Additionally, U.S. Patent 5,925,805 to Ohlrogge *et al.* provides an *Arabidopsis* acetyl CoA carboxylase gene that can be used to increase the oil content of plants. However, the oil content was not increased substantially in these experiments.

A need therefore exists for an improved method to alter the oil content of plants, and in particular to increase the oil content of plants and seeds.

SUMMARY OF INVENTION

The present invention includes materials and provides a method for increasing the oil content of a plant by disrupting the function of a phenylpropanoid pathway protein in that plant. In the present invention, the phenylpropanoid pathway protein is selected from the group

consisting of chalcone synthase (CHS), leucoanthocyanidin dioxygenase (LDOX), phenylalanine ammonia lyase (PAL) and the transcription factor TTG1. In one embodiment, the function of the phenylpropanoid pathway protein is disrupted by suppressing the expression of the gene for that protein. In a further embodiment, the function of the phenylpropanoid pathway protein is disrupted by causing a mutation in the coding or regulatory sequence for that protein.

The present invention also includes and provides a method for increasing the protein content of a plant by disrupting the function of a phenylpropanoid pathway protein in that plant. In the present invention, the phenylpropanoid pathway protein is selected from the group consisting of CHS, PAL, TTG1, and LDOX. In one embodiment, the function of the phenylpropanoid pathway protein is disrupted by suppressing the expression of the gene for that protein. In a further embodiment, the function of the phenylpropanoid pathway protein is disrupted by causing a mutation in the coding sequence for that protein. In another embodiment the increased protein is in the seed of the plant.

The present invention also includes and provides a method for altering the fiber content of a plant by disrupting the function of a phenylpropanoid pathway protein in that plant. In the present invention the phenylpropanoid pathway protein is selected from the group consisting of CHS, PAL, TTG1, and LDOX. In one embodiment the function of the phenylpropanoid pathway protein is disrupted by suppressing the expression of the gene for that protein. In a further embodiment the function of the phenylpropanoid pathway protein is disrupted by causing a mutation in the coding or regulatory sequence for that protein. In another embodiment, the altered fiber content is in the seed of the plant.

The present invention also provides isolated polynucleotides encoding transcription factors that are operative in the phenylpropanoid pathway in a plant. In one embodiment of the present invention, the isolated polynucleotides encoding transcription factors are selected from the group consisting of SEQ ID NOs: 2, 3, and 147, and complements thereof. In a further embodiment, the isolated polynucleotides encoding transcription factors are isolated from corn, canola or soybean.

In another embodiment, the present invention provides vectors comprising at least 50 base pairs from one polynucleotide selected from the group consisting of SEQ ID NOs: 4 through 17, 29, 30, 32 through 96, 128 through 140, 144, 149, 150, and 154 through 165, wherein the polynucleotide can be expressed in both the sense and antisense orientations.

Expression of both sense and antisense orientations can be accomplished by expression of polynucleotides in the sense and antisense orientation in the same or distinct T-DNAs, or by the expression of self-complementary (*i.e.*, hairpin) RNAs. In a further embodiment, the present invention provides expression cassettes comprising said vectors, wherein the polynucleotide of interest is operably linked to a promoter. The invention further includes and provides plants comprising said expression cassettes.

The present invention further provides a method of producing a plant with increased oil content, comprising disrupting the function of a protein in the phenylpropanoid pathway. In a further embodiment of the present invention, the protein in the phenylpropanoid pathway is selected from the group consisting of CHS, PAL, TTG1, and LDOX. In another embodiment, the increased oil content is in the seed of the plant.

The present invention further provides a method of producing a plant with increased protein content, comprising disrupting the function of a protein in the phenylpropanoid pathway. In a further embodiment of the present invention, the protein in the phenylpropanoid pathway is selected from the group consisting of CHS, PAL, TTG1, and LDOX. In another embodiment, the increased protein content is in the seed of the plant.

The present invention further provides a method of increasing phenylpropanoid pathway products in a plant by increasing the expression of phenylpropanoid pathway proteins. In a further embodiment of the present invention the proteins in the phenylpropanoid pathway are selected from the group consisting of CHS, PAL, TTG1, and LDOX.

The invention further provides plants derived from said method and grain derived from said plants. The present invention further provides oil, feed and meal derived from said grain.

DESCRIPTION OF THE FIGURES

Figure 1 shows a schematic representation of the phenylpropanoid pathway in plants.

Figure 2 shows a schematic representation of pMON75702.

Figure 3 shows alignment of predicted protein sequences of LDOX from two wild type *Arabidopsis* ecotypes Columbia (SEQ ID NO: 124) and Landsberg (SEQ ID NO: 125) and the *oil 4* mutant from *Arabidopsis* (SEQ ID NO: 126).

Figure 4 shows a schematic representation of pMON65446.

Figure 5 shows a schematic representation of pMON69847.

Figure 6 shows a schematic representation of pMON82708.

BRIEF DESCRIPTION OF THE SEQUENCES

SEQ ID NO: 1 is a TTG1 protein from *Arabidopsis thaliana*.

SEQ ID NO: 2 is a polynucleotide encoding a TTG1 from soybean.

SEQ ID NO: 3 is a polynucleotide encoding a TTG1 from corn.

SEQ ID NOs: 4 through 8 are polynucleotides encoding CHS from corn.

SEQ ID NOs: 9 through 11 are polynucleotides encoding CHS from soybean.

SEQ ID NOs: 12 and 13 are polynucleotides encoding CHS from corn.

SEQ ID NOs: 14 through 28 are polynucleotides encoding CHS from soybean.

SEQ ID NO: 29 is a polynucleotide encoding CHS from *Brassica napus*.

SEQ ID NO: 30 is a partial polynucleotide sequence encoding CHS from *Brassica napus*.

SEQ ID NO: 31 is a polynucleotide encoding CHS from *Arabidopsis thaliana*.

SEQ ID NO: 32 is a polynucleotide encoding CHS from barley.

SEQ ID NOs: 33 through 36 are polynucleotides encoding CHS from *Brassica napus*.

SEQ ID NOs: 37 through 49 are polynucleotides encoding CHS from corn.

SEQ ID NOs: 50 through 58 are polynucleotides encoding CHS from cotton.

SEQ ID NOs: 59 and 60 are polynucleotides encoding CHS from potato.

SEQ ID NOs: 61 and 62 are polynucleotides encoding CHS from rice.

SEQ ID NOs: 63 and 64 are polynucleotides encoding CHS from sorghum.

SEQ ID NOs: 65 through 78 are polynucleotides encoding CHS from soybean.

SEQ ID NO: 79 is a full length sequence of a polynucleotide encoding CHS from *Brassica napus*.

SEQ ID NOs: 80 and 81 are polynucleotides encoding CHS from soybean.

SEQ ID NOs: 82 and 83 polynucleotides encoding CHS from tomato.

SEQ ID NOs: 84 through 94 are polynucleotides encoding CHS from wheat.

SEQ ID NO: 95 is a polynucleotide encoding a CHS from sunflower.

SEQ ID NO: 96 is a polynucleotide encoding a CHS from tobacco.

SEQ ID NOs: 97 through 120 are primers used for PCR amplification.

SEQ ID NOs: 121 through 123 are polynucleotide sequences encoding LDOX from *Arabidopsis thaliana* cultivars Columbia, Landsberg and an *oil4* mutant, respectively.

SEQ ID NOs: 124 through 126 are the encoded protein sequences of LDOX from *Arabidopsis thaliana* cultivars Columbia, Landsberg and an *oil4* mutant, respectively.

SEQ ID NO: 127 is a polynucleotide encoding LDOX from *Arabidopsis thaliana.*

SEQ ID NO: 128 is a polynucleotide encoding LDOX from barley.

SEQ ID NOs: 129 and 130 are polynucleotides encoding LDOX from *Brassica napus.*

SEQ ID NO: 131 is a polynucleotide encoding LDOX from corn.

SEQ ID NOs: 132 through 134 are polynucleotides encoding LDOX from cotton.

SEQ ID NOs: 135 and 136 are polynucleotides encoding LDOX from onion.

SEQ ID NOs: 137 and 138 polynucleotides encoding LDOX from rice.

SEQ ID NO: 139 is a polynucleotide encoding LDOX from sorghum.

SEQ ID NO: 140 is a polynucleotide encoding LDOX from soybean.

SEQ ID NO: 141 is a polynucleotide encoding a TTG1 from *Arabidopsis thaliana.*

SEQ ID NOs: 142 and 143 are primer sequences used in PCR amplification reactions.

SEQ ID NO: 144 is a polynucleotide encoding a full length sequence of LDOX from *Brassica napus.*

SEQ ID NO: 145 is a polynucleotide encoding LDOX from *Arabidopsis thaliana.*

SEQ ID NO: 146 is a primer sequence used in PCR amplification reactions.

SEQ ID NO: 147 is a polynucleotide encoding TTG1 from *Brassica napus*, contained with the plasmid pMON65420.

SEQ ID NO: 148 is a polynucleotide encoding a CHS from *Arabidopsis thaliana.*

SEQ ID NOs: 149 and 150 are polynucleotides encoding CHS from *Brassica napus.*

SEQ ID NO: 151 is a polynucleotide encoding PAL1 from *Arabidopsis thaliana.*

SEQ ID NO: 152 is a polynucleotide encoding PAL2 from *Arabidopsis thaliana.*

SEQ ID NO: 153 is a polynucleotide encoding PAL3 from *Arabidopsis thaliana.*

SEQ ID NOs: 154 through 160 are polynucleotides encoding PAL from corn.

SEQ ID NOs: 161 through 165 are polynucleotides encoding PAL from rice.

SEQ ID NOs: 166 and 167 are primers used in PCR amplification reactions.

SEQ ID NO: 168 is a CHS protein from *Arabidopsis thaliana*.

SEQ ID NOs: 169 and 170 are primers used in PCR amplification reactions.

SEQ ID NOs: 171 through 176 are primers used in PCR amplification reactions.

DETAILED DESCRIPTION OF THE INVENTION

The present invention provides isolated polynucleotides encoding transcription factors and enzymes important in the phenylpropanoid biosynthetic pathway in a plant. The invention also provides gene constructs that disrupt the function of proteins in the phenylpropanoid biosynthetic pathway. The result of disrupting the function of these proteins in plants is increased oil levels in the plant. Additionally, disrupting the function of these proteins results in increased protein and altered fiber levels in the plant.

Nucleic Acids of Present Invention

A "phenylpropanoid" compound is a compound having a 3-carbon side chain on an aromatic ring, a common structural element shared by all of the metabolites in the monolignol pathway as well as other subclasses of plant phenolics such as flavonoids. A schematic representation of the phenylpropanoid pathway is shown in Figure 1. This figure provides only a condensed overview of a complex pathway and is intended to further illustrate the current invention. It in no way should be construed as a limitation in scope of the described invention.

Flavonoids consist of various groups of compounds, which include chalcones, flavonones, isoflavanones, aurones, catechins, anthocyanidins, and flavonols. These plant metabolites influence many plant-animal interactions and also function to protect plants against UV-B irradiation. The first committed step of flavonoid biosynthesis from the phenylpropanoid pathway is catalyzed by the enzyme chalcone synthase (CHS). Three molecules of acetate-derived malonyl-CoA and one molecule of p-coumaryl-CoA are condensed to generate one molecule of tetrahydroxychalcone. Details of the phenylpropanoid pathway can be found in most plant biochemistry references (*see* for example, *Biochemistry and Molecular Biology of Plants*, Buchanan, Gruissem, and Jones. eds., American Society of Plant Physiology (2000)).

In one embodiment, the present invention provides constructs that disrupt the function of phenylpropanoid pathway enzymes such as CHS, PAL and LDOX. In further embodiments, the instant invention provides constructs that disrupt the function of transcription factors, such as TTG1 (TRANSPARENT TESTA GLABRA1, *see*, Walker *et al.*, *The Plant Cell,* 11:1337-49 (1999)), which regulate the transcription of phenylpropanoid pathway genes.

In some embodiments, the polynucleotides of the present invention are cloned, amplified, or otherwise constructed from plants generally regarded as oilseed crops. In preferred embodiments, the plants are soybean, canola or corn.

The present invention includes a plurality of polynucleotides that encode for the identical amino acid sequence. The degeneracy of the genetic code allows for several trinucleotide codons to encode the same amino acid. Additionally, the present invention includes isolated nucleic acids comprising allelic variants. The term “allelic” as used herein refers to a polynucleotide that encodes an alternative amino acid sequence that has the same or similar function of the original form of the same gene.

The nucleic acids and vectors of the present invention need not have the exact nucleic acid sequences described herein. Instead, the sequences of the nucleic acids and vectors can vary, so long as the nucleic acid either performs the function for which it is intended or has some other utility, for example as a nucleic acid probe for complementary nucleic acids. For example, some sequence variability in any part of the TTG1 nucleic acid sequence is permitted so long as the variant retains at least 10% of the activity observed under similar conditions for analogous wild-type transcription factors.

Fragment and variant nucleic acids are also encompassed by this invention. Nucleic acid fragments encompassed by the invention are of two types. First, fragment nucleic acids that are not full length but do perform their intended function are encompassed within this invention. Second, fragments of the nucleic acids identified herein are useful as hybridization probes or in one of the gene suppression strategies useful in down regulation of enzyme activity.

"Substantially similar" nucleic acid fragments refers to fragments wherein changes in one or more nucleotide bases result in substitution of one or more amino acids but do not substantially affect the functional properties of the polypeptide encoded by the nucleotide sequence. "Substantially similar" also refers to nucleic acid fragments wherein changes in one or more nucleotide bases do not affect the ability of the nucleic acid fragment to mediate alteration of gene expression by gene suppression strategies useful in down regulation of

enzyme activity. "Substantially similar" also refers to modifications of the nucleic acid fragments of the instant invention, such as deletion or insertion of one or more nucleotides that do not substantially affect the functional properties of the resulting transcript vis-à-vis the ability to mediate gene silencing or alteration of the functional properties of the resulting protein molecule. It is therefore understood that the invention encompasses more than the specific exemplary nucleotide or amino acid sequences and includes functional equivalents thereof.

Substantially similar nucleic acid fragments may be selected by screening nucleic acid fragments representing subfragments or modifications of the nucleic acid fragments of the present invention, wherein one or more nucleotides are substituted, deleted or inserted, for their ability to affect the level of the polypeptide encoded by the unmodified nucleic acid fragment in a plant or plant cell. For example, a substantially similar nucleic acid fragment comprising a fragment with a change of at least 1 of 25 contiguous nucleotides derived from the instant nucleic acid fragment can be constructed and introduced into a plant or plant cell. The level of the polypeptide encoded by the unmodified nucleic acid fragment present in a plant or plant cell exposed to the substantially similar nucleic fragment can then be compared to the level of the polypeptide in a plant or plant cell that is not exposed to the substantially similar nucleic acid fragment.

For example, it is well known in the art that antisense suppression and co-suppression of gene expression may be accomplished using nucleic acid fragments representing less than the entire coding region of a gene, untranslated region, or intron and by nucleic acid fragments that do not share 100% the entire coding region of a gene, untranslated region, or intron and by nucleic acid fragments that do not share 100% sequence identity with the gene to be suppressed. Moreover, alterations in a nucleic acid fragment that result in the production of a chemically equivalent amino acid at a given site, but do not affect the functional properties of the encoded polypeptide, are well known in the art.

The present invention also provides methods for detection and isolation of derivative or variant nucleic acids encoding the proteins provided herein. The methods involve hybridizing at least a portion of a nucleic acid comprising any part of SEQ ID NOs: 4 through 17, 29, 30, 32 through 96, 128 through 140, 144, 149, 150, and 154 through 165, to a sample nucleic acid, thereby forming a hybridization complex, and detecting the hybridization complex. The presence of the complex correlates with the presence of a derivative or variant nucleic acid that can be further characterized by nucleic acid sequencing, expression of RNA or protein and

testing to determine whether the derivative or variant retains activity. In general, the portion of a nucleic acid comprising any part of SEQ ID NOs: 4 through 17, 29, 30, 32 through 96, 128 through 140, 144, 149, 150, and 154 through 165, used for hybridization is at least 15 nucleotides, and hybridization is under hybridization conditions that are sufficiently stringent to permit detection and isolation of substantially homologous nucleic acids. Such conditions are well known to one of skill in the art. Exemplary high stringency conditions include hybridization in 50% formamide, 5X SSC, 50mM sodium phosphate, pH 7.0, 5mM EDTA, 0.1% SDS, 5X Denhardt's with 100 μg/mL denatured salmon sperm DNA at 42°C, and a wash in 0.1X SSC, 0.1% SDS at 60 to 65°C. An extensive guide to the hybridization of nucleic acids is found in Tijssen, *Laboratory Techniques in Biochemistry and Molecular Biology-Hybridization with Nucleic Acid Probes*, Part 1, Chapter 2 (Elsevier, New York: 1993); Ausubel *et al.*, eds., *Current Protocols in Molecular Biology*, Chapter 2 (Greene Publishing and Wiley - Interscience, New York: 1995). *See* Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual* (2d ed., Cold Spring Harbor Laboratory Press, Plainview, New York: 1989). Using these references, those of ordinary skill can generate variants of the present nucleic acids.

Computer analyses can also be utilized for comparison of sequences to determine sequence identity. Such analyses include, but are not limited to, CLUSTAL in the PC/Gene program (available from Intelligenetics, Mountain View, California); the ALIGN program (Version 2.0) and GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Version 8 (available from Genetics Computer Group (GCG), 575 Science Drive, Madison, Wisconsin). Alignments using these programs can be performed using the default parameters. The CLUSTAL program is well described by Higgins *et al.*, *Gene*, 73:237-244 (1988); Higgins *et al.*, *CABIOS*, 5:151-153 (1989); Corpet *et al.*, *Nucleic Acids Res.*, 16:10881-10890 (1988); Huang *et al.*, *CABIOS*, 8:155-165 (1992); and Pearson *et al.*, *Meth. Mol. Biol.*, 24:307-331 (1994). The BLAST programs of Altschul *et al.*, *J. Mol. Biol.*, 215:403 (1990), are based on the algorithm of Karlin and Altschul, *Proc. Natl. Acad. Sci. (U.S.A.)*, 87:2264-2268 (1990). To obtain gapped alignments for comparison purposes, Gapped BLAST (in BLAST 2.0) can be utilized as described in Altschul *et al.*, *Nucleic Acids Res.*, 25:3389 (1997). Alternatively, PSI-BLAST (in BLAST 2.0) can be used to perform an iterated search that detects distant relationships between molecules. When utilizing BLAST, Gapped BLAST, PSI-BLAST, the default parameters of the respective programs can be used. The BLASTN program (for nucleotide sequences) uses as defaults a word length (W) of 11, an expectation (E) of 10, a cutoff of 100, M=5, N=-4, and a comparison of both strands. For amino acid sequences,

the BLASTP program uses as defaults a word length (W) of 3, an expectation (E) of 10, and the BLOSUM62 scoring matrix (*see*, Henikoff & Henikoff, *Proc. Natl. Acad. Sci. (U.S.A.)*, 89:10915 (1989)). Alignment may also be performed manually by inspection.

For purposes of the present invention, comparison of nucleotide sequences for determination of percent sequence identity to the nucleic acid sequences disclosed herein is preferably made using the BLASTN program (version 1.4.7 or later) with its default parameters or any equivalent program. By "equivalent program" is intended any sequence comparison program that, for any two sequences in question, generates an alignment having identical nucleotide or amino acid residue matches and an identical percent sequence identity when compared to the corresponding alignment generated by the preferred program. In the present invention, it is anticipated that the nucleotides herein could be useful at about 70%, about 80%, about 90%, about 95%, or about 99% identity, in addition to the recited usefulness at 100% identity.

Isolation of Nucleic Acids Encoding Phenylpropanoid Pathway Proteins

Nucleic acids encoding phenylpropanoid pathway proteins can be identified and isolated by standard methods as described by Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor, NY (1989). For example, a DNA sequence encoding a target protein can be identified by screening a genomic DNA or cDNA library generated from nucleic acid derived from a particular cell type, cell line, primary cells, or tissue.

Screening for DNA fragments that encode a target phenylpropanoid pathway protein and regulators can be accomplished by screening plaques from a genomic or cDNA library for hybridization to a probe of an available phenylpropanoid pathway protein from other organisms or by screening plaques from a cDNA expression library for binding to antibodies that specifically recognize the target phenylpropanoid pathway protein. DNA fragments that hybridize to phenylpropanoid pathway protein probes from other organisms or plaques carrying DNA fragments that are immunoreactive with antibodies to phenylpropanoid pathway proteins can be subcloned into a vector and sequenced or used as probes to identify other cDNA or genomic sequences encoding all or a portion of the desired phenylpropanoid pathway protein gene.

A cDNA library can be prepared, for example, by random oligo priming or oligo dT priming. Plaques containing DNA fragments can be screened with probes or antibodies specific for phenylpropanoid pathway proteins. DNA fragments encoding a portion of the target gene

can be subcloned and sequenced and used as probes to identify a genomic copy of the gene. DNA fragments encoding a portion of the target protein can be verified by determining sequence homology with other known genes encoding the target protein or by hybridization to messenger RNA specific to the target protein. Once cDNA fragments encoding portions of the 5', middle and 3' ends of a target protein are obtained, they can be used as probes to identify and clone a complete genomic copy of the gene for the target protein from a genomic library.

Portions of the genomic copy or copies of a phenylpropanoid pathway protein gene can be sequenced and the 5' end of the gene identified by standard methods including either DNA sequence homology to analogous genes or by RNAase protection analysis, as described by Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Press, Cold Spring Harbor, New York (1989). The 3' and 5' ends of the target gene can also be located by computer searches of genomic sequence databases using known coding regions specific to the target protein. Once portions of the 5' end of the gene are identified, complete copies of the target gene can be obtained by standard methods, including cloning or polymerase chain reaction (PCR) synthesis using oligonucleotide primers complementary to the DNA sequence at the 5' end of the gene. The presence of an isolated full-length copy of the target phenylpropanoid pathway gene can be verified by hybridization, partial sequence analysis, or by expression of the protein.

A further embodiment of the present invention is the isolation from an EST library of nucleic acids encoding proteins in the phenylpropanoid pathway. EST libraries may be prepared as described in Cooke *et al.*, "EST and genomic sequencing projects", In *Plant Gene Isolation: Principles and Practice*, ed. Foster and Twell, pp. 401-419: John Wiley & Sons LTD (1996), and references included therein.

Gene Suppression

Gene suppression may be effective against a native plant gene associated with a trait, *e.g.,* to provide plants with reduced levels of a protein encoded by the native gene or with enhanced or reduced levels of an affected metabolite.

As used herein "gene suppression" means any of the well-known methods for suppressing a transcript or a protein from a gene including post-transcriptional gene suppression and transcriptional suppression. Post transcriptional gene suppression is mediated by transcribed RNA having homology to a gene targeted for suppression. The RNA transcribed from the suppressing transgene can be in the sense orientation to effect what is called

co-suppression, in the anti-sense orientation to effect what is called anti-sense suppression or in both orientations producing a double-stranded RNA to effect what is called RNA interference (RNAi). Transcriptional suppression is mediated by a transcribed double-stranded RNA having homology to promoter DNA sequence to effect what is called promoter *trans* suppression.

More particularly, post-transcriptional gene suppression by anti-sense oriented RNA to regulate gene expression in plant cells is disclosed in U.S. Patent 5,107,065 (Shewmaker *et al.*) and U.S. Patent 5,759,829 (Shewmaker *et al.*). Post-transcriptional gene suppression by sense-oriented RNA to regulate gene expression in plants is disclosed in U.S. Patent 5,283,184 (Jorgensen *et al.*) and U.S. Patent 5,231,020 (Jorgensen *et al.*). Post transcriptional gene suppression by double-stranded RNA to suppress genes in plants by RNAi is disclosed in PCT Publication WO 99/53050 (Waterhouse *et al.*) using recombinant DNA constructs comprising sense-oriented and anti-sense-oriented elements of a targeted gene in separate transcription units or in a single transcription unit. *See*, also PCT Publication WO 99/49029 (Graham *et al.*), U.S. Publication 2003/0175965 A1 (Lowe *et al.*), U.S. Patent Application 10/465,800 (Fillatti), and U.S. Patent 6,506,559 (Fire *et al.*). *See*, also Titia Sijen *et al.*, *The Plant Cell*, 8:2277-2294, December 1996, which discloses the use of constructs carrying inverted repeats of a cowpea mosaic virus gene in transgenic plants to mediate virus resistance. Another DNA construct for RNAi gene suppression comprising a singly-oriented gene element bordered by oppositely-oriented promoters is disclosed in U.S. Publication 2003/0061626 A1 (Plaetinck *et al.*) and U.S. Patent 6,326,193. *See*, also U.S. Patent Application 10/393,347, which discloses constructs and methods for simultaneously expressing one or more recombinant genes while simultaneously suppressing one or more native genes in a transgenic plant. See also U.S. Patent 6,448,473, which discloses multigene expression vectors for use in plants. All of the above-described patents, applications and international publications disclosing materials and methods for post-transcriptional gene suppression in plants are incorporated herein by reference.

A preferred method of post transcriptional gene suppression in plants employs either sense-oriented or anti-sense-oriented, transcribed RNA that is stabilized, *e.g.*, with a terminal hairpin structure. A preferred DNA construct for effecting post transcriptional gene suppression is transcribed to a segment of anti-sense oriented RNA having homology to a gene targeted for suppression, where the anti-sense RNA segment is followed at the 3' end by a contiguous, complementary, shorter segment of RNA in the sense orientation. The use of self-stabilized anti-sense RNA oligonucleotides in plants is disclosed in PCT Publication

WO 94/01550 (Agrawal *et al.*). See also PCT Publication WO 98/05770 (Werner *et al.*), where the anti-sense RNA is stabilized by hairpin forming repeats of poly(CG) nucleotides. See also U.S. Publication 2002/0048814 A1 (Oeller), where sense or anti-sense RNA is stabilized by a poly(T)-poly(A) tail. See also U.S. Publication 2003/0018993 A1 (Gutterson *et al.*), where sense or anti-sense RNA is stabilized by an inverted repeat of a subsequence of an NOS gene. See also U.S. Publication 2003/0036197 A1 (Glassman *et al.*), where RNA having homology to a target is stabilized by two complementary RNA regions. All of the above-described patents, applications and international publications disclosing materials and methods for employing stabilized RNA and its use in gene suppression in plants are incorporated herein by reference.

Transcriptional suppression such as promoter *trans* suppression can be effected by a expressing a DNA construct comprising a promoter operably linked to inverted repeats of promoter DNA for a target gene. Constructs useful for such gene suppression mediated by promoter *trans* suppression are disclosed by Mette *et al., The EMBO Journal*, 18(1):241-148 (1999), and by Mette *et al*., *The EMBO Journal,* 19(19):5194-5201 (2000), both of which are incorporated herein by reference.

Yet another strategy for gene down-regulation or suppression contemplated by this invention is the use of ribozymes. Ribozymes, or catalytic RNA molecules capable of cleaving target mRNA at specific sites, are well know in the art. (*See*, for example, Gibson and Shillitoe, *Molecular Biotechnology,* 7(2):125-137 (1997)).

Another method for abolishing the expression of a gene is by insertion mutagenesis using the T-DNA of *Agrobacterium tumefaciens*. After generating the insertion mutants, the mutants can be screened to identify those containing the insertion in a phenylpropanoid pathway protein gene. Plants containing a single transgene insertion event at the desired gene can be crossed to generate homozygous plants for the mutation (Koncz *et al.,* In: *Methods in Arabidopsis Research*, World Scientific (1992)).

Alternatively, a plant phenotype can be altered by eliminating phenylpropanoid pathway protein genes, such as CHS, PAL, TTG1, and LDOX, e.g., by homologous recombination (Kempin *et al., Nature,* 389:802 (1997)).

A plant trait can also be modified by using the cre-lox system (U.S. Patent 5,658,772). A plant genome can be modified to include first and second lox sites that are then contacted with a Cre recombinase. If the lox sites are in the same orientation, the intervening DNA

sequence between the two sites is excised. If the lox sites are in the opposite orientation, the intervening sequence is inverted.

One of ordinary skill in the art will recognize that a number of methods can be used to inactivate, abolish or suppress gene expression or the activity of the phenylpropanoid pathway proteins of the present invention.

Expression Cassettes and Vectors

An "expression cassette" refers to a nucleic acid construct that when introduced into a host cell, results in transcription or translation of an RNA or polypeptide, respectively. Antisense or sense constructs that are not or cannot be translated are expressly included by this definition. An expression cassette may be a combination of nucleic acid elements that perform specific functions and are maintained as a clonable unit.

A "vector " refers to a nucleic acid molecule capable of carrying inserted nucleic acid sequences and being perpetuated in a host cell.

The expression cassettes and vectors of the invention include nucleic acids encoding phenylpropanoid pathway proteins, in sense and antisense configuration. Additionally, this invention contemplates vectors and cassettes constructed to disrupt the function of proteins in the phenylpropanoid biosynthetic pathway. In addition, this invention sets forth the expression cassettes and vectors for overexpressing novel transcription factor genes (such as TTG1) in plants to increase products in the phenylpropanoid pathway.

A transgene comprising a phenylpropanoid pathway protein can be subcloned into an expression cassette or vector, and its expression can be detected or quantified. This method of screening is useful to identify transgenes providing for an expression of a phenylpropanoid pathway protein, and expression of a phenylpropanoid pathway protein in a transformed plant cell.

Plasmid vectors that provide for easy selection, amplification, and transformation of the transgene in prokaryotic and eukaryotic cells include, for example, pUC-derived vectors, pSK-derived vectors, pGEM-derived vectors, pSP-derived vectors, pBS-derived vectors, pFastBac (Invitrogen) for baculovirus expression and pYES2 (Invitrogen) for yeast expression. Additional elements may be present in such vectors, including origins of replication to provide for autonomous replication of the vector, selectable marker genes, preferably encoding antibiotic or herbicide resistance, unique multiple cloning sites providing for multiple sites to insert DNA sequences or genes encoded in the transgene, and sequences that enhance

transformation of prokaryotic and eukaryotic cells. One vector that is useful for expression in both plant and prokaryotic cells is the binary Ti plasmid (Schilperoort *et al.*, U.S. Patent 4,940,838) as exemplified by vector pGA582. This binary Ti plasmid vector has been previously characterized by An, *Methods in Enzymology,* 153:292 (1987). This binary Ti vector can be replicated in bacteria, such as *E. coli* and *Agrobacterium.* The *Agrobacterium* plasmid vectors can also be used to transfer the transgene to plant cells. The binary Ti vectors preferably include the nopaline T-DNA right and left borders to provide for efficient plant cell transformation, a selectable marker gene, unique multiple cloning sites in the T border regions, the *col*E1 replication of origin and a wide host range replicon. The binary Ti vectors carrying a transgene of the invention can be used to transform both prokaryotic and eukaryotic cells but are preferably used to transform plant cells. *See*, for example, Glassman *et al.*, U.S. Patent 5,258,300.

A vector or expression cassette of the present invention can contain a transgene encoding a phenylpropanoid pathway protein in sense or antisense configuration. The vectors and expression cassettes of the present invention are constructed according to whether the strategy is to overexpress or suppress the function of the phenylpropanoid pathway protein. The term "expression", as used herein, refers to the transcription and stable accumulation of sense (mRNA) or antisense RNA derived from the nucleic acid fragment of the invention. Expression may also refer to translation of mRNA into a polypeptide.

In general, the expression vectors and cassettes of the invention contain at least a promoter capable of expressing RNA in a plant cell and a terminator, in addition to a nucleic acid encoding a phenylpropanoid pathway protein. Other elements may also be present in the expression cassettes of the invention. For example, expression cassettes can also contain enhancers, introns, untranslated leader sequences, cloning sites, matrix attachment regions for silencing the effects of chromosomal control elements, and other elements known to one of skill in the art.

Nucleic acids encoding phenylpropanoid pathway proteins are operably linked to regulatory elements such as promoter, termination signals and the like. Operably linking a nucleic acid under the regulatory control of a promoter or a regulatory element means positioning the nucleic acid such that the expression of the nucleic acid is controlled by these sequences. In general, promoters are found positioned 5' (upstream) to the nucleic acid that that they control. Thus, in the construction of heterologous promoter/nucleic acid combinations, the promoter is preferably positioned upstream to the nucleic acid and at a distance from the

transcription start site of the nucleic acid that the distance between the promoter and the transcription start site approximates the distance observed in the natural setting. As is known in the art, some variation in this distance can be tolerated without loss of promoter function. Similarly, the preferred positioning of a regulatory element with respect to a heterologous nucleic acid placed under its control is the natural position of the regulatory element relative to the structural gene it naturally regulates. Again, as is known in the art, some variation in this distance can be accommodated.

Expression cassettes have promoters that can regulate gene expression. Promoter regions are typically found in the flanking DNA sequence upstream from coding regions in both prokaryotic and eukaryotic cells. A promoter sequence provides for regulation of transcription of the downstream gene sequence and typically includes from about 50 to about 2,000 nucleotide base pairs. Promoter sequences also contain regulatory sequences, such as enhancer sequences, that can influence the level of gene expression. Some isolated promoter sequences can provide for gene expression of heterologous genes, that is, a gene different from the native or homologous gene. Promoter sequences are also known to be strong or weak or inducible. A strong promoter provides for a high level of gene expression, whereas a weak promoter provides for a very low level of gene expression. An inducible promoter is a promoter that provides for turning on and off of gene expression in response to an exogenously added agent or to an environmental or developmental stimulus. Promoters can also provide for tissue-specific or developmental regulation. An isolated promoter sequence that is a strong promoter for heterologous genes is advantageous because it provides for a sufficient level of gene expression to allow for easy detection and selection of transformed cells and provides for a high level of gene expression when desired. Transcription initiation regions that are preferentially expressed in seed tissue and that are undetectable in other plant parts are considered desirable for seed oil modifications in order to minimize any disruptive or adverse effects of the gene product.

Promoters of the instant invention will generally include, but are not limited to, promoters that function in plant cells. Useful promoters include the napin promoter (Kridl *et al.*, *Seed Sci. Res.*, 1:209-219 (1991)), the WEREWOLF promoter (Lee and Schiefelbein, *Cell*, 99:473-483 (1999)), the e35S CaMV promoter for the 35S transcript RNA derived from Cauliflower Mosaic Virus modified as described by Kay *et al.*, *Science*, 235:1299–1302 (1987), the 35S promoter (Odell *et al.*, *Nature*, 313:810 (1985)), the CaMV 19S (Lawton *et al.*, *Plant Mol. Biol.*, 9, 31F (1987)), *nos* (Ebert *et al.*, *Proc. Nat. Acad. Sci. (U.S.A.)*, 84:5745 (1987)), Adh (Walker *et al.*, *Proc. Nat. Acad. Sci. (U.S.A.)*, 84:6624 (1987)), sucrose synthase (Yang *et*

al., Proc. Nat. Acad. Sci. (U.S.A.), 87:4144 (1990)), tubulin, actin (Wang *et al., Mol. Cell. Biol.*, 12:3399 (1992)), *cab* (Sullivan *et al., Mol. Gen. Genet.*, 215:431 (1989)), PEPCase promoter (Hudspeth *et al., Plant Mol. Biol.*, 12:579 (1989)), the 7S-alpha'-conglycinin promoter (Beachy *et al., EMBO J.,* 4:3047 (1985)), or those associated with the R gene complex (Chandler *et al., The Plant Cell,* 1:1175 (1989)). Other useful promoters include the tomato E8, patatin, ubiquitin, mannopine synthase (mas), soybean seed protein glycinin (Gly), soybean vegetative storage protein (vsp) promoters, and the *Arabidopsis* banyuls promoter.

Indeed, in an embodiment, the promoter used is a seed coat--specific promoter. Examples of seed coat--regulated genes and transcriptional regions are disclosed herein. Other promoters known to function, for example, in maize include the promoters for the following genes: *waxy, Brittle, Shrunken 2*, Branching enzymes I and II, starch synthases, debranching enzymes, oleosins, glutelins, sucrose synthases, and globulin1. Other promoters useful in the practice of the invention that are known to those of skill in the art are also contemplated by the invention.

Moreover, transcription enhancers or duplications of enhancers can be used to increase expression from a particular promoter. Examples of such enhancers include, but are not limited to, elements from the CaMV 35S promoter and octopine synthase genes (Last *et al.*, U.S. Patent 5,290,924). As the DNA sequence between the transcription initiation site and the start of the coding sequence, *i.e.,* the untranslated leader sequence, can influence gene expression, one may also wish to employ a particular leader sequence. Any leader sequence available to one of skill in the art may be employed. Preferred leader sequences direct optimum levels of expression of the attached gene, for example, by increasing or maintaining mRNA stability or by preventing inappropriate initiation of translation. The choice of such sequences is at the discretion of those of skill in the art. Sequences that are derived from genes that are highly expressed in higher plants, and in soybean, corn, and canola in particular, are contemplated.

Expression cassettes of the present invention will also include a sequence near the 3' end of the cassette that acts as a signal to terminate transcription from a heterologous nucleic acid and that directs polyadenylation of the resultant mRNA. Some 3' elements that can act as termination signals include those from the nopaline synthase gene of *Agrobacterium tumefaciens* (Bevan *et al., Nucl. Acid Res.*, 11:369 (1983)), a napin 3' untranslated region, a globulin 3' untranslated region, the tumor large locus (tml) from *Agrobacterium tumefaciens* (Barker *et al., Plant Mol. Biol.*, 2:335-350 (1983)), or one from a zein gene, such as, for

example Z27. Other 3' elements known to those of skill in the art also can be used in the vectors of the present invention.

Regulatory elements, such as Adh intron 1 (Callis *et al., Genes Develop.*, 1:1183 (1987)), a rice actin intron (McElroy *et al., Mol. Gen. Genet.*, 231 (1):150-160 (1991)), sucrose synthase intron (Vasil *et al., Plant Physiol.*, 91:5175 (1989)), the maize HSP70 intron or TMV omega element (Gallie *et al., The Plant Cell,* 1:301 (1989)), may further be included where desired. Other such regulatory elements useful in the practice of the present invention are known to those of skill in the art and can also be placed in the vectors of the present invention.

The vectors of the present invention can be optimized for expression in plants by having one or more codons replaced by other codons encoding the same amino acids so that the polypeptide is optimally translated by the translation machinery of the plant species in which the vector is used. *See,* for example Perlak *et al., Proc. Natl. Acad. Sci. (U.S.A.),* 88(8):3324-3328 (1991)) or Moriyama and Powell, *Journal of Molecular Evolution,* 45(5):514-523 (1997).

Plant Transformation

In a preferred embodiment of the invention, a transgenic plant expressing the desired protein is to be produced. Various methods for the introduction of a desired polynucleotide sequence encoding the desired protein into plant cells are available and known to those of skill in the art and include, but are not limited to, (1) physical methods such as microinjection (Capecchi, *Cell,* 22(2):479-488 (1980)), electroporation (Fromm *et al., Proc. Natl. Acad. Sci. (U.S.A.),* 82(17):5824-5828 (1985); U.S. Patent No. 5,384,253), and microprojectile mediated delivery (biolistics or gene gun technology) (Christou *et al., Bio/Technology,* 9:957 (1991); Fynan *et al., Proc. Natl. Acad. Sci. (U.S.A.),* 90(24):11478-11482 (1993)); (2) virus mediated delivery methods (Clapp, *Clin. Perinatol.*, 20(1):155-168 (1993); Lu *et al., J. Exp. Med.,* 178(6):2089-2096 (1993); Eglitis and Anderson, *Biotechniques*, 6(7):608-614 (1988)); and (3) *Agrobacterium*-mediated transformation methods.

The most commonly used methods for transformation of plant cells are the *Agrobacterium*-mediated DNA transfer process (Fraley *et al., Proc. Natl. Acad. Sci. (U.S.A.),* 80:4803 (1983)) and the biolistics or microprojectile bombardment mediated process (*i.e.,* the gene gun). Typically, nuclear transformation is desired but where it is desirable to specifically transform plastids, such as chloroplasts or amyloplasts, plant plastids may be transformed

utilizing a microprojectile-mediated delivery of the desired polynucleotide for certain plant species such as tobacco, *Arabidopsis*, potato, and *Brassica* species.

Agrobacterium-mediated transformation is achieved through the use of a genetically engineered soil bacterium belonging to the genus *Agrobacterium*. Several *Agrobacterium* species mediate the transfer of a specific DNA known as "T-DNA", which can be genetically engineered to carry any desired piece of DNA into many plant species. The major events marking the process of T-DNA mediated pathogenesis are induction of virulence genes, processing and transfer of T-DNA. This process is the subject of many reviews (Ream, *Ann. Rev. Phytopathol.*, 27:583-618 (1989); Howard and Citovsky, *Bioassays*, 12:103-108 (1990); Kado, *Crit. Rev. Plant Sci.*, 10:1-32 (1991); Zambryski, *Annual Rev. Plant Physiol. Plant Mol. Biol.*, 43:465-490 (1992); Gelvin, In *Transgenic Plants*, Kung and Wu, eds., Academic Press, San Diego, California, pp. 49-87 (1993); Binns and Howitz, In *Bacterial Pathogenesis of Plants and Animals* (Dang, ed.) Berlin: Springer Verlag, pp. 119-138 (1994); Hooykaas and Beijersbergen, *Ann. Rev. Phytopathol.*, 32:157-179 (1994); Lessl and Lanka, *Cell,* 77:321-324 (1994); Zupan and Zambryski, *Annual Rev. Phytopathol.,* 27:583-618 (1995)).

Agrobacterium-mediated genetic transformation of plants involves several steps. The first step, in which the virulent *Agrobacterium* and plant cells are first brought into contact with each other, is generally called "inoculation". The *Agrobacterium*-containing solution is then removed from contact with the explant by draining or aspiration. Following the inoculation, the *Agrobacterium* and plant cells/tissues are permitted to be grown together for a period of several hours to several days or more under conditions suitable for growth and T-DNA transfer. This step is termed "co-culture". Following co-culture and T-DNA delivery, the plant cells are treated with bactericidal or bacteriostatic agents to prevent further growth of the *Agrobacterium* remaining in contact with the explant or in the vessel containing the explant. If this is done in the absence of any selective agents to promote preferential growth of transgenic versus non-transgenic plant cells, then this is typically referred to as the "delay" step. If done in the presence of selective pressure favoring transgenic plant cells, then it is referred to as a "selection" step. When a "delay" is used, it is typically followed by one or more "selection" steps. Both the "delay" and "selection" steps typically include bactericidal or bacteriostatic agents to prevent further growth of any remaining *Agrobacterium* cells because the growth of *Agrobacterium* cells is undesirable after the infection (inoculation and co-culture) process.

A number of wild-type and disarmed strains of *Agrobacterium tumefaciens* and *Agrobacterium rhizogenes* harboring Ti or Ri plasmids can be used for gene transfer into plants.

The *Agrobacterium* hosts contain disarmed Ti and Ri plasmids that do not contain the oncogenes that cause tumorigenesis or rhizogenesis, respectfully, which are used as the vectors and contain the genes of interest that are subsequently introduced into plants. Preferred strains would include but are not limited to *Agrobacterium tumefaciens* strain C58, a nopaline-type strain that is used to mediate the transfer of DNA into a plant cell, octopine-type strains such as LBA4404 or succinamopine-type strains, *e.g.*, EHA101 or EHA105. The nucleic acid molecule, prepared as a DNA composition *in vitro*, is introduced into a suitable host such as *E. coli* and mated into the *Agrobacterium*, or directly transformed into competent *Agrobacterium*. These techniques are well-known to those of skill in the art.

The *Agrobacterium* can be prepared either by inoculating a liquid such as Luria Burtani (LB) media directly from a glycerol stock or streaking the *Agrobacterium* onto a solidified media from a glycerol stock, allowing the bacteria to grow under the appropriate selective conditions, generally from about 26°C-30°C, and taking a single colony or a small loop of *Agrobacterium* from the plate and inoculating a liquid culture medium containing the selective agents. Those of skill in the art are familiar with procedures for growth and suitable culture conditions for *Agrobacterium* as well as subsequent inoculation procedures. The density of the *Agrobacterium* culture used for inoculation and the ratio of *Agrobacterium* cells to explant can vary from one system to the next, and therefore optimization of these parameters for any transformation method is expected.

Typically, an *Agrobacterium* culture is inoculated from a streaked plate or glycerol stock and is grown overnight and the bacterial cells are washed and resuspended in a culture medium suitable for inoculation of the explant.

With respect to microprojectile bombardment (U.S. Patents 5,550,318; 5,538,880; and 5,610,042; and PCT Publication WO 95/06128; each of which is specifically incorporated herein by reference in its entirety), particles are coated with nucleic acids and delivered into cells by a propelling force. Exemplary particles include those comprised of tungsten, platinum, and preferably, gold. It is contemplated that in some instances DNA precipitation onto metal particles would not be necessary for DNA delivery to a recipient cell using microprojectile bombardment. However, it is contemplated that particles may contain DNA rather than be coated with DNA. Hence, it is proposed that DNA-coated particles may increase the level of DNA delivery via particle bombardment but are not, in and of themselves, necessary.

For the bombardment, cells in suspension are concentrated on filters or solid culture medium. Alternatively, immature embryos or other target cells may be arranged on solid culture medium. The cells to be bombarded are positioned at an appropriate distance below the macroprojectile stopping plate.

An illustrative embodiment of a method for delivering DNA into plant cells by acceleration is the Biolistics Particle Delivery System (BioRad, Hercules, California), which can be used to propel particles coated with DNA or cells through a screen, such as a stainless steel or Nytex screen, onto a filter surface covered with plant cells cultured in suspension. The screen disperses the particles so that they are not delivered to the recipient cells in large aggregates. It is believed that a screen intervening between the projectile apparatus and the cells to be bombarded reduces the size of projectiles aggregate and may contribute to a higher frequency of transformation by reducing the damage inflicted on the recipient cells by projectiles that are too large.

For microprojectile bombardment, one will attach (*i.e.*, "coat") DNA to the microprojectiles such that it is delivered to recipient cells in a form suitable for transformation thereof. In this respect, at least some of the transforming DNA must be available to the target cell for transformation to occur, while at the same time during delivery the DNA must be attached to the microprojectile. Therefore, availability of the transforming DNA from the microprojectile may comprise the physical reversal of bonds between transforming DNA and the microprojectile following delivery of the microprojectile to the target cell. This need not be the case, however, as availability to a target cell may occur as a result of breakage of unbound segments of DNA or of other molecules which comprise the physical attachment to the microprojectile. Availability may further occur as a result of breakage of bonds between the transforming DNA and other molecules, which are either directly or indirectly attached to the microprojectile. It further is contemplated that transformation of a target cell may occur by way of direct recombination between the transforming DNA and the genomic DNA of the recipient cell. Therefore, as used herein, a "coated" microprojectile will be one that is capable of being used to transform a target cell, in that the transforming DNA will be delivered to the target cell, yet will be accessible to the target cell such that transformation may occur.

Any technique for coating microprojectiles that allows for delivery of transforming DNA to the target cells may be used. Methods for coating microprojectiles that have been demonstrated to work well with the current invention have been specifically disclosed herein. DNA may be bound to microprojectile particles using alternative techniques, however. For

example, particles may be coated with streptavidin and DNA end labeled with long chain thiol cleavable biotinylated nucleotide chains. The DNA adheres to the particles due to the streptavidin-biotin interaction but is released in the cell by reduction of the thiol linkage through reducing agents present in the cell.

Alternatively, particles may be prepared by functionalizing the surface of a gold oxide particle, providing free amine groups. DNA, having a strong negative charge, binds to the functionalized particles. Furthermore, charged particles may be deposited in controlled arrays on the surface of mylar flyer disks used in the PDS-1000 Biolistics device, thereby facilitating controlled distribution of particles delivered to target tissue.

As disclosed above, it further is proposed that the concentration of DNA used to coat microprojectiles may influence the recovery of transformants containing a single copy of the transgene. For example, a lower concentration of DNA may not necessarily change the efficiency of the transformation but may instead increase the proportion of single copy insertion events. In this regard, approximately 1 ng to 2000 ng of transforming DNA may be used per each 1.8 mg of starting microprojectiles.

Microprojectile bombardment techniques are widely applicable and may be used to transform virtually any plant species. Examples of species that have been transformed by microprojectile bombardment include monocot species such as maize (PCT Publication WO 95/06128), barley, wheat (U.S. Patent 5,563,055, specifically incorporated herein by reference in its entirety), rice, oat, rye, sugarcane, and sorghum; as well as a number of dicots including tobacco, soybean (U.S. Patent 5,322,783, specifically incorporated herein by reference in its entirety), sunflower, peanut, cotton, tomato, and legumes in general (U.S. Patent 5,563,055, specifically incorporated herein by reference in its entirety).

For microprojectile bombardment transformation in accordance with the current invention, both physical and biological parameters may be optimized. Physical factors are those that involve manipulating the DNA/microprojectile precipitate or those that affect the flight and velocity of either the macro- or microprojectiles. Biological factors include all steps involved in manipulation of cells before and immediately after bombardment, such as the osmotic adjustment of target cells to help alleviate the trauma associated with bombardment, the orientation of an immature embryo or other target tissue relative to the particle trajectory, and also the nature of the transforming DNA, such as linearized DNA or intact supercoiled

plasmids. It is believed that pre-bombardment manipulations are especially important for successful transformation of immature embryos.

Accordingly, it is contemplated that one may wish to adjust various of the bombardment parameters in small scale studies to fully optimize the conditions. One may particularly wish to adjust physical parameters such as DNA concentration, gap distance, flight distance, tissue distance, and helium pressure. It further is contemplated that the grade of helium may effect transformation efficiency. One also may optimize the trauma reduction factors (TRFs) by modifying conditions which influence the physiological state of the recipient cells and which may therefore influence transformation and integration efficiencies. For example, the osmotic state, tissue hydration and the subculture stage or cell cycle of the recipient cells may be adjusted for optimum transformation.

To select or score for transformed plant cells regardless of transformation methodology, the DNA introduced into the cell contains a gene that functions in a regenerable plant tissue to produce a compound that confers upon the plant tissue resistance to an otherwise toxic compound. Genes of interest for use as a selectable, screenable, or scorable marker would include but are not limited to GUS, green fluorescent protein (GFP), luciferase (LUX), antibiotic or herbicide tolerance genes. Examples of antibiotic resistance genes include the penicillins, kanamycin (and neomycin, G418, bleomycin); methotrexate (and trimethoprim); chloramphenicol; kanamycin and tetracycline.

Particularly preferred selectable marker genes for use in the present invention would include genes that confer resistance to compounds such as antibiotics like kanamycin (*nptII*), hygromycin B (*aph IV*) and gentamycin (*aac3* and *aac*C4) (Dekeyser *et al.*, *Plant Physiol.,* 90:217-223 (1989)), and herbicides like glyphosate (Della-Cioppa *et al.*, *Bio/Technology*, 5:579-584 (1987)). Other selection devices can also be implemented including but not limited to tolerance to phosphinothricin, bialaphos, and positive selection mechanisms (Joersbo *et al.*, *Mol. Breed.,* 4:111-117 (1998)) and are considered within the scope of the present invention.

The regeneration, development, and cultivation of plants from various transformed explants is well documented in the art. This regeneration and growth process typically includes the steps of selecting transformed cells and culturing those individualized cells through the usual stages of embryonic development through the rooted plantlet stage. Transgenic embryos and seeds are similarly regenerated. The resulting transgenic rooted shoots are thereafter planted in an appropriate plant growth medium such as soil. Cells that survive the exposure to

the selective agent, or cells that have been scored positive in a screening assay, may be cultured in media that supports regeneration of plants. In an exemplary embodiment, MS and N6 media may be modified by including further substances such as growth regulators. A preferred growth regulator for such purposes is dicamba or 2,4-D. However, other growth regulators may be employed, including NAA, NAA + 2,4-D or perhaps even picloram. Media improvement in these and like ways has been found to facilitate the growth of cells at specific developmental stages. Tissue may be maintained on a basic media with growth regulators until sufficient tissue is available to begin plant regeneration efforts, or following repeated rounds of manual selection, until the morphology of the tissue is suitable for regeneration, at least 2 weeks, then transferred to media conducive to maturation of embryoids. Cultures are transferred every 2 weeks on this medium. Shoot development will signal the time to transfer to medium lacking growth regulators.

The transformed cells, identified by selection or screening and cultured in an appropriate medium that supports regeneration, will then be allowed to mature into plants. Developing plantlets are transferred to soiless plant growth mix, and hardened off, *e.g.*, in an environmentally controlled chamber at about 85% relative humidity, 600 ppm CO_2, and 25-250 microeinsteins m^{-2} s^{-1} of light, prior to transfer to a greenhouse or growth chamber for maturation. Plants are preferably matured either in a growth chamber or greenhouse. Plants are regenerated from about 6 wk to 10 months after a transformant is identified, depending on the initial tissue. During regeneration, cells are grown on solid media in tissue culture vessels. Illustrative embodiments of such vessels are petri dishes and Plant Cons. Regenerating plants are preferably grown at about 19 to 28°C. After the regenerating plants have reached the stage of shoot and root development, they may be transferred to a greenhouse for further growth and testing.

Note, however, that seeds on transformed plants may occasionally require embryo rescue due to cessation of seed development and premature senescence of plants. To rescue developing embryos, they are excised from surface-disinfected seeds 10-20 days post-pollination and cultured. An embodiment of media used for culture at this stage comprises MS salts, 2% sucrose, and 5.5 g/L agarose. In embryo rescue, large embryos (defined as greater than 3 mm in length) are germinated directly on an appropriate media. Embryos smaller than that may be cultured for 1 wk on media containing the above ingredients along with 10^{-5}M abscisic acid and then transferred to growth regulator-free medium for germination.

The present invention can be used with any transformable cell or tissue. By transformable as used herein is meant a cell or tissue that is capable of further propagation to give rise to a plant. Those of skill in the art recognize that a number of plant cells or tissues are transformable in which after insertion of exogenous DNA and appropriate culture conditions the plant cells or tissues can form into a differentiated plant. Tissue suitable for these purposes can include but is not limited to immature embryos, scutellar tissue, suspension cell cultures, immature inflorescence, shoot meristem, nodal explants, callus tissue, hypocotyl tissue, cotyledons, roots, and leaves.

Any suitable plant culture medium can be used. Examples of suitable media would include but are not limited to MS-based media (Murashige and Skoog, *Physiol. Plant*, 15:473-497 (1962)) or N6-based media (Chu *et al.*, *Scientia Sinica*, 18:659 (1975)) supplemented with additional plant growth regulators including but not limited to auxins such as picloram (4-amino-3,5,6-trichloropicolinic acid), 2,4-D (2,4-dichlorophenoxyacetic acid) and dicamba (3,6-dichloroanisic acid); cytokinins such as BAP (6-benzylaminopurine) and kinetin; ABA; and gibberellins. Other media additives can include but are not limited to amino acids, macroelements, iron, microelements, vitamins and organics, carbohydrates, undefined media components such as casein hydrolysates, with or without an appropriate gelling agent such as a form of agar, such as a low melting point agarose or Gelrite if desired. Those of skill in the art are familiar with the variety of tissue culture media, which when supplemented appropriately, support plant tissue growth and development and are suitable for plant transformation and regeneration. These tissue culture media can either be purchased as a commercial preparation, or custom prepared and modified. Examples of such media would include but are not limited to Murashige and Skoog (Murashige and Skoog, *Physiol. Plant,* 15:473-497 (1962)), N6 (Chu *et al., Scientia Sinica,* 18:659 (1975)), Linsmaier and Skoog (Linsmaier and Skoog, *Physio. Plant.,* 18:100 (1965)), Uchimiya and Murashige (Uchimiya and Murashige, *Plant Physiol.,* 15:473 (1962)), Gamborg's B5 media (Gamborg *et al.*, *Exp. Cell Res.,* 50:151 (1968)), D medium (Duncan *et al.*, *Planta,* 165:322-332 (1985)), McCown's Woody plant media (McCown and Lloyd, *HortScience,* 16:453 (1981)), Nitsch and Nitsch (Nitsch and Nitsch, *Science,* 163:85-87 (1969)), and Schenk and Hildebrandt (Schenk and Hildebrandt, *Can. J. Bot.,* 50:199-204 (1972)) or derivations of these media supplemented accordingly. Those of skill in the art are aware that media and media supplements such as nutrients and growth regulators for use in transformation and regeneration and other culture

conditions such as light intensity during incubation, pH, and incubation temperatures can be optimized for the particular variety of interest.

Any of the nucleic acid molecules of the invention may be introduced into a plant cell in a permanent or transient manner in combination with other genetic elements such as vectors, promoters, enhancers, and the like. Furthermore, any of the nucleic acid molecules of the invention may be introduced into a plant cell in a manner that allows for expression or overexpression of the protein or fragment thereof encoded by the nucleic acid molecule.

It is also to be understood that two different transgenic plants can also be mated to produce offspring that contain two independently segregating added, exogenous genes. Selfing of appropriate progeny can produce plants that are homozygous for both added, exogenous genes that encode a polypeptide of interest. Backcrossing to a parental plant and out-crossing with a non-transgenic plant are also contemplated, as is vegetative propagation.

Transgenic plants may find use in the commercial manufacture of proteins or other molecules, where the molecule of interest is extracted or purified from plant parts, seeds, and the like. Cells or tissue from the plants may also be cultured, grown *in vitro*, or fermented to manufacture such molecules.

The transgenic plants may also be used in commercial breeding programs or may be crossed or bred to plants of related crop species. Improvements encoded by the recombinant DNA may be transferred, *e.g.,* from cells of one species to cells of other species, *e.g.,* by protoplast fusion.

Plants

Plants for use with the vectors of the invention include dicots and monocots, preferably oil producing species including but not limited to, corn (*Zea mays*), *Brassica* sp., particularly those *Brassica* species useful as sources of seed oil (*e.g., B. napus, B. rapa, B. juncea*), sunflower (*Helianthus annuus*), safflower (*Carthamus tinctorius*), rice (*Oryza sativa*), soybean (*Glycine max*), tobacco (*Nicotiana tabacum*), peanuts (*Arachis hypogaea*), cotton (*Gossypium barbadense, Gossypium hirsutum*), coconut (*Cocos nucifera*), cocoa (*Theobroma cacao*), oil palm (*Elaeis guineensis*), flax (*Linum usitatissimum*); *Cuphea* species; castor (*Ricinus communis*), olive (*Olea europaea*), cashew (*Anacardium occidentale*), macadamia (*Macadamia integrifolia*), and almond (*Prunus amygdalus*)).

The plants described in the present invention can be used to generate seeds containing valuable commercial products, such as oil and animal feed. The processing and milling of such

seed is well known in the art (*see*, for example, Watson in *Corn and Corn Improvement*, Holt, Nelson, and Keeney, eds., American Society of Agronomy pp. 881-940 (1988); and *Soybeans: Improvement, Production and Uses*, Wilcox, ed. (1987)).

The present invention will be further described by reference to the following detailed examples, which are not to be viewed as limiting the scope of the invention as further described herein. It is understood that there are many extensions, variations, and modifications on the basic theme of the present invention beyond that shown in the examples and description, which are within the spirit and scope of the present invention.

Example 1

This example sets forth the identification and sequence analysis of a chalcone synthase ortholog from *Brassica napus* and TTG1 orthologs from corn, soybean and canola.

The *Arabidopsis thaliana* chalcone synthase (CHS) nucleic acid sequence (GenBank Accession M20308, SEQ ID NO: 31) was used as a BLAST query against proprietary *Brassica napus* sequence libraries. A single EST clone from a sequence library that was derived from whole seeds at 30 days after pollination, LIB4156-020-R1-K1-E4 (SEQ ID NO: 149), was identified as having high homology. A 944-base pair fragment containing the CHS coding sequence was removed from LIB4156-020-R1-K1-E4 by digestion with SmaI and XbaI. The fragment was purified and sequenced using methodology well known in the art. This purified fragment containing the *Brassica napus* CHS ortholog (BnCHS) (SEQ ID NO: 33) was then cloned in the antisense orientation between the e35S promoter and the tml 3' UTR in a backbone vector. The resulting plasmid, containing the e35S::BnCHS::tml 3' UTR, was named KAHADF032096 and was subsequently used in the construction of plant transformation vectors.

The *Arabidopsis thaliana* TTG1 nucleic acid sequence (GenBank Accession AJ133743, SEQ ID NO: 141) was used as a BLAST query against proprietary *Brassica napus* sequence libraries. A single EST clone that was derived from silique walls at 10 days after pollination, LIB3169-025-P1-K2-B5 (SEQ ID NO: 79), was identified as having high homology. This sequence was used as a query against all *Arabidopsis* DNA sequences in GenBank®, the NIH genetic sequence database containing an annotated collection of all publicly available DNA sequences (*Nucleic Acids Research* 30(1):17-20 (2002)). The BLAST search identified *Arabidopsis thaliana* TTG1 as having the highest homology, thus confirming the *Brassica* sequence as a TTG1 ortholog.

The coding sequence for *Brassica napus* TTG1 (BnTTG1) was PCR amplified from the EST clone LIB3169-025-P1-K2-B5 (SEQ ID NO: 30) using the primers ttg-3' (SEQ ID NO: 142) and ttg-5' (SEQ ID NO: 143).

These primers were designed to add SalI and PstI restriction sites as well as a Kozak consensus sequence to the TTG1 sequence to facilitate further cloning. Sequencing was done using the protocol and equipment supplied by ABI PRISM (ABI PRISM BigDye Terminators (v 3.0) Ready Reaction Cycle Sequencing Kit and the ABI PRISM 377 Automated DNA Sequencer, Foster City, CA). Alignments with the predicted *Arabidopsis* TTG1 peptide revealed that clone LIB3169-025-P1-K2-B5 started at base pair +3 of the predicted coding sequence. This was the reason that a Kozak consensus translation initiation site was also added by the primers described above. The resulting PCR product was isolated, purified, and cloned into pCR2.1TOPO, purchased from Invitrogen (Carlsbad, CA), following the manufacturer's instructions. The resulting plasmid containing the full-length coding sequence of TTG1 was named pMON65420 and was subsequently used in the construction of plant transformation vectors.

Example 2

This example describes the construction of plant transformation vectors with the BnCHS and BnTTG1 coding sequences described above.

<u>pMON75051 (e35S::BnCHS(antisense)::tml)</u>

A 3470-base pair fragment containing the e35S promoter, the BnCHS (SEQ ID NO: 33) in antisense orientation and the tml 3' UTR was removed from the vector KAHADF032096 by digestion with NotI. The fragment was ligated into the vector pCGN11121, which had been digested with NotI. The vector pCGN11121 contains a nopaline T-DNA right border sequence and octopine T-DNA left border sequence, in between which are the promoter from the Figwort Mosaic Virus (FMV) and the intron from *Z. mays* HSP70 driving the expression of the CP4 EPSP synthase gene containing a CTP, linked to a synthetic EPSP synthase coding region and the 3' UTR from the pea rbsc E9 gene and the recognition sites for the recombinase. The resulting plasmid was named pMON75051. The nucleic acid sequence was determined using standard sequencing methodology and confirmed the integrity of the cloning junctions.

<u>pMON75052 (napin::BnCHS(antisense)::napin)</u>

A 1280-base pair fragment containing BnCHS (SEQ ID NO: 33) coding sequence was removed from LIB4156-020-R1-K1-E4 by digestion with NotI and PstI. The fragment was

isolated and purified using methodology well known in the art and then cloned in the antisense orientation between the napin promoter and napin 3' UTR in pMON67163 which had been digested with NotI-Sse8387I. The resulting plasmid was named pMON75052. The nucleic acid sequence was determined using standard sequencing methodology and confirmed the integrity of the cloning junctions.

pMON75053 (pBAN::BnCHS(antisense)::napin)

The NotI/PstI fragment described above was also cloned in the antisense orientation down stream of the banyuls promoter (pBAN) by ligating it into a NotI/Sse387I digested pMON70905 and upstream of the napin 3' UTR, for transformation into canola. The resulting plasmid was named pMON75053, and is notated as pBAN::BnCHS(antisense)::napin.

pMON75702 (35S::BnTTG1(antisense)::tml)

The *Brassica napus* TTG1 ortholog (BnTTG1) from pMON65419 is cloned, in the antisense orientation, downstream of the e35S promoter and upstream of the tml 3' UTR, for transformation into canola. The resulting plasmid is named pMON75702 (Figure 2).

pMON70906 (pBAN::BnTTG1(antisense)::napin)

The vector pMON67163 contains a nopaline T-DNA right border sequence and an octopine T-DNA left border sequence. Between the two T-DNA border sequences are the 35S promoter driving expression of an CP4 EPSP synthase coding sequence consisting of the first exon of the *Arabidopsis* EPSP synthase gene (containing a chloroplast targeting sequence) linked to a synthetic EPSP synthase coding region and the 3' UTR from the pea rbcS E9 gene; the napin promoter and 3' UTR and recognition signals for cre recombinase.

The napin promoter was removed from pMON67163 by sequential digestion with PacI followed by NotI. The PacI overhang was blunt ended using Pfu polymerase according to the manufacturer's instructions (Stratagene, La Jolla, CA). The resulting 9503-base pair fragment was ligated to a 994-base pair BamHI-NotI fragment from pMON69809, which contained the *Arabidopsis* banyuls promoter (pBAN). The BamHI overhang was blunt ended using Pfu polymerase according to the manufacturer's instructions (Stratagene). The resulting plasmid was named pMON70905. The nucleic acid sequence was determined using standard sequencing methodology and confirmed the integrity of the cloning junctions.

The vector pMON65419 contains EST clone LIB3169-025-P1-K2-B5 in a pSPORT1 (Invitrogen) background. A 1362-base pair NotI-PstI fragment containing the coding sequence

of BnTTG1 (SEQ ID NO: 79) was ligated into a 10477-base pair NotI-Sse8387I fragment from pMON70905. The resulting plasmid, which contains the BnTTG1 coding sequence in antisense orientation being driven by the pBAN promoter, was named pMON70906. The nucleic acid sequence was determined using standard sequencing methodology (ABI PRISM BigDye Terminators (v 3.0) Ready Reaction Cycle Sequencing Kit and the ABI PRISM 377 Automated DNA Sequencer, (Foster City, CA) and confirmed the integrity of the cloning junctions.

pMON65423 (e35S::BnTTG1(sense)::tml)

The TTG1 coding sequence of pMON65420 (SEQ ID NO: 147) was found to contain a single nucleotide change (C273T), relative to that shown in SEQ ID NO: 30, that does not result in a change in the predicted protein sequence. A 1031-base pair fragment containing the BnTTG1 coding sequence (SEQ ID NO: 147) was removed from pMON65420 by digestion with SalI and PstI. The resulting fragment was ligated into the vector pCGN9977, which had been sequentially digested with SalI and PstI. The vector pCGN9977 contains a nopaline T-DNA right border sequence and an octopine T-DNA left border. Between the two T-DNA border sequences are (a) the e35S promoter, (b) a polylinker to facilitate cloning, (c) the 3' untranslated region of the tumor large locus from *Agrobacterium tumefaciens* (tml), (d) the 35S promoter, driving expression of the bar gene from *Streptomyces hygroscopicus*, encoding phosphinothricin acetyl transferase (EMBL:X17220), and (e) the 3' untranslated region of the nopaline synthase gene of *Agrobacterium tumefaciens* (nos). The resulting plasmid, containing the BnTTG1 coding sequence in sense orientation between the e35S promoter and the tml 3' UTR, was named pMON65423. This vector was designed for complementation studies in *Arabidopsis* ttg1 mutants. The nucleic acid sequence was determined using standard sequencing methodology and confirmed the integrity of the cloning junctions.

pMON65440 (pBAN::BnTTG1(sense)::napin)

The coding region of the *Brassica napus* TTG1 ortholog was removed from pMON65423, as above, by digestion with NcoI and PstI. The DNA fragment is ligated between the pBAN promoter and napin 3' UTR in NcoI-Sse8387I digested pMON70905. The resulting plasmid, pMON65440, contains the BnTTG1 coding sequence in sense orientation being driven by the pBAN promoter and using the napin 3' UTR. This vector was designed for complementation studies in *Arabidopsis* ttg1 mutants.

pMON65442 (pWER::BnTTG1(sense)::tml)

An *E. coli* culture containing the BAC clone T9L3 was obtained from the Arabidopsis Biological Resource Center (ABRC, Columbus, OH). The BAC clone DNA was isolated by standard methods known to one of skill in the art.

The promoter of the *Arabidopsis* WEREWOLF gene (Lee and Schiefelbein, *Cell* 99:473-483 (1999)) was isolated by PCR from BAC clone T9L3 using the primers WER_Nco (SEQ ID NO: 166) and were 5' #2 (SEQ ID NO: 167).

The reaction conditions for the PCR reaction followed a protocol recommended by the enzyme manufacturer (PE Applied Biosystems, Foster City, CA). pWEREWOLF was amplified using 1.5 µL of T9L3 DNA as template, 30 nanomoles each of the primers WER_Nco (SEQ ID NO: 166) and were 5' #2 (SEQ ID NO: 167), 10 micromoles each of dATP, dCTP, dGTP and TTP, 2.5 units of AmpliTaq Gold (PE Applied Biosystems) in 1X Opti-Prime™ Buffer 3 (Stratagene). After an initial incubation at 95°C for 10 minutes, 28 cycles of PCR were performed with 94°C for 15 seconds, 62°C for 10 seconds, 52°C for 10 seconds, and 72°C for 3 minutes followed by 1 cycle of 72°C for 7 minutes. The product of the PCR reaction was purified according to standard methodology well known in the art and cloned into pCR2.1 Topo (Invitrogen) according to manufacturer's instructions. The resulting plasmid was named pMON69835. The sequence of this clone was determined using standard sequencing methodologies as set forth by PE Applied Biosystems (Foster City, CA).

The Banyuls promoter was removed from pMON65423 by digestion with PvuII and NcoI. The resulting 10966-bp fragment was ligated to a 1098-bp fragment from pMON69835, which contained the *Arabidopsis* WEREWOLF promoter. The resulting plasmid contained the BnTTG1 coding sequence in sense orientation between the WEREWOLF promoter and the tml 3'UTR. This vector was designed for complementation studies in *Arabidopsis* ttg1 mutants.

Table 1 summarizes the vectors described above.

Table 1. Summary of Constructs for Plant Transformations

Plasmid	Promoter Coding	Sequence Orientation	3' UTR	Crop
pMON65423	e35S	BnTTG1 Sense	tml	*Arabidopsis*
pMON65440	pBAN	BnTTG1 Sense	napin	*Arabidopsis*
pMON65442	pWEREWOLF	BnTTG1 Sense	tml	*Arabidopsis*
pMON70906	pBAN	BnTTG1 Antisense	napin	Canola
pMON75702	e35S	BnTTG1 Antisense	tml	Canola
pMON75051	e35S	BnCHS Antisense	tml	Canola
pMON75052	napin	BnCHS Antisense	napin	Canola
pMON75053	pBAN	BnCHS Antisense	napin	Canola

Example 3

This example describes the transformation of canola and *Arabidopsis* plants.

***Arabidopsis* Transformation**

Arabidopsis seeds, mutant and wild type, are sown onto 2 ¼ inch pots containing reverse osmosis water (ROW) saturated MetroMix 200 (The Scotts Company, Columbus, OH). The plants are vernalized by placing the pots in a flat, covered with a humidity dome, in a growth chamber at 4-7°C, 8-hrs light/day for 4-7 days. The flats are transferred to a growth chamber at 22°C, 55% relative humidity, and 16-hrs light/day at an average intensity of 160-200 μEinstein/s/m^2. After germination, the dome is lifted and slid back about 1 inch to allow for mild air circulation without desiccation. The humidity dome is removed once true leaves have formed. The plants are bottom watered, as needed, with ROW until well established, generally 2-3 weeks after germination. Plants are then bottom watered, as needed, with Plantex 15-15-18 solution at 50 ppm N_2. Pots are thinned so that 1 plant remains per pot at 2-3 weeks after germination. Once plants begin to bolt, the primary inflorescence is trimmed to encourage the growth of axillary bolts.

The vectors are introduced into *Agrobacterium tumefaciens* strain ABI using methodologies well known in the art.

Transgenic *Arabidopsis thaliana* plants are obtained essentially as described by Bent *et al.* (*Science* 265:1856-1860 (1994)) or Bechtold *et al.* (*C.R. Acad. Sci., Life Sciences*, 316:1194-1199 (1993)). Cultures containing the desired vector are grown overnight in Luria Broth (10% bacto-tryptone, 5% yeast extract, and 10% NaCl) with kanamycin (75mg/L), chloramphenicol (25mg/L), and spectinomycin (100mg/L). The bacterial culture is centrifuged

and resuspended in 5% sucrose + .05% Silwet-77. The aerial portions of whole *Arabidopsis thaliana* (CS89; ttg1) plants (at ~5-7 weeks of age) are immersed in the resulting solution for 2-3 seconds. The excess solution is removed by blotting the plants on paper towels. The dipped plants are placed on their side in a covered flat and transferred to a growth chamber at 19°C. After 16 to 24 hours, the dome is removed and the plants are set upright. When plants have reached maturity, water is withheld for 2-7 days prior to seed harvest. Harvested seed is passed through a stainless steel mesh screen (40 holes/inch) to remove debris. Seed is stored in paper coin envelopes at room temperature.

Seeds are sown onto flats containing ROW saturated MetroMix 200. The plants are vernalized and germinated as described above. After true leaves have emerged, the aerial portions of the seedlings are sprayed with a solution containing a 1:200 dilution of Finale herbicide (The Scotts Company) in ROW. Approximately 1 week after the first application, the plants are sprayed a second time. Up to 16 Finale resistant seedlings are transplanted to a 2 ¼" pot, one seedling per pot, containing MetroMix 200 and are grown under the conditions described above.

Alternatively, transformed plants are selected on media. *Arabidopsis* seeds are surface sterilized using a vapor phase sterilization protocol. An open container of seeds is placed in a dessicator with a beaker containing 100 mL of household bleach. Immediately prior to sealing the dessicator, 3 mL of concentrated HCl is added to the bleach. The dessicator is sealed and a vacuum is applied to allow sterilization by chlorine fumes. Seeds are incubated for several hours. Sterilized seed are sprinkled onto *Arabidopsis* Germination Media [MS Salts (1X); sucrose (1%); myo-Inositol (100mg/L); Thiamine-HCl (1mg/L); Pyridoxine-HCl (50 mg/L); Nicotinic Acid (500 mg/L); MES pH 5.7 (0.05%) and Phytagar (0.7%) supplemented with 25mg/L glufosinate ammonium (Sigma)].

Canola Transformation

Seeds of *Brassica napus* cv Ebony are planted in 2-inch pots containing Metro Mix 350 (The Scotts Company). The plants are grown in a growth chamber at 24°C, and a 16/8 hour photoperiod, with light intensity of 400 μEinstein/s/m^2 (HID lamps). After 2-1/2 weeks, the plants are transplanted into 6-inch pots and grown in a growth chamber at 15/10°C day/night temperature, 16/8 hour photoperiod, light intensity of 800 μEinstein/s/m^2 (HID lamps).

Four terminal internodes from plants just prior to bolting or in the process of bolting but before flowering are removed and surface sterilized in 70% v/v ethanol for 1 minute, 2% w/v

sodium hypochlorite for 20 minutes and rinsing 3 times with sterile deionized water. Six to seven stem segments are cut into 5mm discs, maintaining orientation of basal end.

The *Agrobacterium* culture is grown overnight on a rotator shaker at 24°C in 2 mL of Luria Broth (10% bacto-tryptone, 5% yeast extract, and 10% NaCl) containing 50 mg/L kanamycin, 24 mg/L chloramphenicol and 100 mg/L spectinomycin. A 1:10 dilution is made in MS media (Murashige and Skoog, *Physiol. Plant*, 15:473-497 (1962)) giving approximately 9×10^8 cells per mL. The stem discs (explants) are inoculated with 1.0 mL of *Agrobacterium* and the excess is aspirated from the explants.

The explants are placed basal side down in petri plates containing media comprising 1/10 MS salts, B5 vitamins, 3% sucrose, 0.8% agar, pH 5.7, 1.0 mg/L 6-benzyladenine (BA). The plates are layered with 1.5 mL of media containing MS salts, B5 vitamins, 3% sucrose, pH 5.7, 4.0 mg/L p-chlorophenoxyacetic acid, 0.005mg/L kinetin and covered with sterile filter paper.

Following a 2- to 3-day co-culture, the explants are transferred to deep dish petri plates containing MS salts, B5 vitamins, 3% sucrose, 0.8% agar, pH 5.7, 1mg/L 6-benzyladenine, 500 mg/L carbenicillin, 50 mg/L cefotaxime, 200 mg/L kanamycin or 175 mg/L gentamycin for selection. Seven explants are placed on each plate. After 3 weeks they are transferred to fresh media, 5 explants per plate. The explants are cultured in a growth room at 25°C, continuous light (Cool White).

Example 4

This example describes the analysis of transformed plants for oil and protein levels in the seed.

Oil levels in *Arabidopsis* seed tissues are established by near-infrared reflectance (NIR) spectroscopy (Williams and Norris (eds.), *Near-infrared Technology in the Agricultural and Food Industries*, American Association of Cereal Chemists, Inc., St. Paul (1987)), whereby NIR spectra of pooled seed samples harvested from individual plants are measured, and oil levels are calculated based on regression analysis using a standard curve generated from analysis of *Arabidopsis* seed with varying oil levels as determined gravimetrically following accelerated solvent extraction (*Better Solutions for Food and Beverage Analysis, 2nd edition,* Dionex Corporation, Sunnyvale, CA 1997).

Oil levels in canola seed are established by NIR as described above using a standard curve generated from analysis of canola seed with varying oil levels. Briefly, mature canola seeds previously dried to less than 10% moisture are equilibrated to ambient humidity in paper envelopes at room temperature. Single replicate sub-samples (2-3 g) are placed in NIR ring cups (aluminum/quartz; 2-inch diameter by 0.5-inch thick; purchased from Foss North America Inc., Silver Springs Maryland), and sealed with a paperboard disk. The loaded ring cups are placed in an autoloader and scanned sequentially on a Foss Analytical model 6500 Spectrometer (Foss North America Inc, Silver Springs, Maryland). Each sample is scanned 25 times from 400 to 2500 nm (resolution 2 nm) and the average spectrum is compiled. The averaged spectrum is reduced to second derivative spectra, smoothed and transformed to a series of principal component scores. The total oil and protein levels are predicted based on previously prepared calibration models. WinISI software (WinISI ver 1.00, Infrasoft International LLC, State College, PA) is used for calibration development and instrument operation.

Oil levels in soybean are established either by NIR as described above, using a standard curve generated from analysis of soybean seed with varying oil levels, or by near-infrared transmittance (NIT). When NIT is employed, the NIT spectra of pooled seed samples harvested from individual plants are measured, and oil levels are calculated based on regression analysis using a standard curve generated from analysis of soybean seed with varying oil levels as determined gravimetrically following accelerated solvent extraction.

Protein Analyses

For seed protein analysis, small bulk samples consisting of 50-100 seeds for each treatment are measured using near infrared reflectance spectroscopy (InfraTec model 1221, Teccator, Hogannas, Sweden). This procedure is based upon the observation that a linear relationship exists between the absorption of near infrared radiation and the quantity of chemical constituents comprised in a typical grain sample. Prior to analyzing unknown samples, spectral data is collected with calibration samples that are subsequently analyzed using a nitrogen combustion analysis technique (Murray and Williams, "Chemical Principles of Near-Infrared Technology", in *Near-Infrared Technology in the Agricultural and Food Industries*, Williams and Norris (eds.), 1987). A multivariate model is developed using the spectral data from the spectrometer and the primary data. In the present case, a PLS-1 (Partial Least Squares Regression Type I) multivariate model is constructed using 152 calibration samples. Each unknown sample is scanned on the spectrometer at least 5 times and its protein

content predicted with each scan. Each time the sample is scanned, it is added back to the sample cuvette to minimize multiplicative scattering effects, which are not correlated to chemical property of interest. The predicted protein is averaged for the multiple scans and then reported for each sample.

Example 5

This example sets forth the identification and sequencing of the mutant *oil4* gene in an *Arabidopsis thaliana*, Landsberg *erecta* (Ler) mutant having high oil, high sum of oil plus protein, and altered seed color phenotype.

Mutagenized (M_2) seeds of *Arabidopsis thaliana,* ecotype Landsberg are obtained both by purchase from Lehle Seeds (Round Rock, Texas, U.S.A.) and by standard EMS mutagenesis methodology. The M_2 plants are grown from the M_2 seeds in greenhouse conditions with one plant per 2.5 inch pot. The resulting M_3 seeds are collected from individual M_2 plants.

Seeds from approximately 5000 M_3 lines of *Arabidopsis thaliana* are analyzed for total oil and protein levels as described above in Example 4, and compared to wild type control plants. A mutant (*oil4*) is identified that has increased levels of oil (39% vs. 32.9% for wild type) and higher oil plus protein (64.6% vs. 60.4% for wild type).

Using map-based cloning techniques (*see*, for example, Jander *et al.*, *Plant Physiol.*, 129(2):440-450 (2002)), the mutant *oil4* gene is mapped to a gene identified as F7H19.60. This gene possesses homology to known leucoanthocyanidin dioxygenases. Based on this homology, F7H19.60 is determined to contain the mutation responsible for the high oil, high sum of oil plus protein and seed color phenotypes in the *oil4* mutant. The sequences of the F7H19.60 gene locus in the wild types and *oil4* mutant are PCR amplified, and determined by standard sequencing methodology. The gene locus, in each case, is amplified using the following sequencing primers. Primer Pair F7H19.60_1 is Forward Primer (SEQ ID NO: 97) and Reverse Primer (SEQ ID NO: 98). Primer Pair F7H19.60_2 is Forward Primer (SEQ ID NO: 99) and Reverse Primer (SEQ ID NO: 100). Primer Pair F7H19.60_3 is Forward Primer (SEQ ID NO: 101) and Reverse Primer (SEQ ID NO: 102). Primer Pair F7H19.60_4 is Forward Primer (SEQ ID NO: 103) and Reverse Primer (SEQ ID NO: 104). Primer Pair F7H19.60_5 is Forward Primer (SEQ ID NO: 105) and Reverse Primer (SEQ ID NO: 106). Primer Pair F7H19.60_6 is Forward Primer (SEQ ID NO: 107) and Reverse Primer (SEQ ID NO: 108). Primer Pair F7H19.60_7 is Forward Primer (SEQ ID NO: 109) and Reverse Primer (SEQ ID NO: 110). Primer Pair F7H19.60_8 is Forward Primer (SEQ ID NO: 111) and

Reverse Primer (SEQ ID NO: 112). Primer Pair F7H19.60_9 is Forward Primer (SEQ ID NO: 113) and Reverse Primer (SEQ ID NO: 114). Primer Pair F7H19.60_10 is Forward Primer (SEQ ID NO: 115) and Reverse Primer (SEQ ID NO: 116). Primer Pair F7H19.60_11 is Forward Primer (SEQ ID NO: 117) and Reverse Primer (SEQ ID NO: 118). Primer Pair F7H19.60_12 is Forward Primer (SEQ ID NO: 119) and Reverse Primer (SEQ ID NO: 120).

The following Polymerase Chain Reaction (PCR) mixture is prepared for each primer pair (10 µL 10X Opti-Prime™ Buffer 3 (Stratagene, La Jolla, CA); 2 µL 20mM dNTPs; 5 µL Template DNA 30; 1.0 µL Taq Gold ; 4 µL F/R Primers (at 20 µM/L) ; 84.0 µL dH_2O). The PCR amplification is carried out using the following Thermocycler program (1. 94°C for 10 minutes; 2. 92°C for 15 seconds; 3. 56°C for 15 seconds; 4. 72°C for 1 minute, 30 seconds; 5. Repeat Steps 2 through 4 an additional 44 times; 6. 72°C for 10 minutes; Hold at 4°C).

The resulting PCR products are purified using the QIAquick PCR Purification Kit (Qiagen, Valencia, CA). The purified PCR products are sequenced using the protocol set forth in the Applied Biosystems (ABI) (Foster City, CA) sequencing protocol and equipment (ABI PRISM BigDye Terminators (v 3.0) Ready Reaction Cycle Sequencing Kit and the ABI PRISM 377 Automated DNA Sequencer).

The LDOX gene (At4g22880; F7H19.60; gi:3292813) from the *oil4* mutant has a nucleotide change that leads to a predicted amino acid substitution of serine for proline at position 148 (Pro148Ser). Figure 3 shows the alignment of the predicted protein sequences of the LDOX gene from the *oil4* mutant (SEQ ID NO: 126) with that of wild type *Arabidopsis* Columbia (Col) (SEQ ID NO: 124) and Landsberg (Ler) (SEQ ID NO: 125) ecotypes. The nucleotide sequence of the LDOX gene for Landsberg and Columbia ecotypes as well as the *oil4* mutant are shown in SEQ ID NOs: 121 through 123.

<u>pMON65435 (p35S::AtLDOX (sense)::tml)</u>

The vector pMON73273 contains a nopaline T-DNA right border sequence and an octopine T-DNA left border sequence. Between the two T-DNA border sequences are the e35S-CMV promoter, a polylinker to facilitate cloning, the 3' untranslated region of the tumor large locus (tml) from *Agrobacterium tumefaciens*, the 35S promoter from the Figwort Mosaic Virus, (FMV), driving expression of an CP4 EPSP synthase coding sequence consisting of the first exon of the *Arabidopsis* EPSP synthase gene (containing a chloroplast targeting sequence) linked to a synthetic EPSP synthase coding region and the 3' untranslated region from the pea rbcS E9 gene; and recognition sites for cre recombinase.

The EST clone LIB3177-088-P1-K1-D7 (SEQ ID NO: 145) is derived from 2-day whole seedlings/seed coats from *Arabidopsis thaliana* ecotype Columbia. This clone, encoding a full length cDNA from locus At4g22880, designated AtLDOX, is removed by digestion with SalI and NotI from a pSPORT1 vector. The resulting approximately 1.3 kb fragment is ligated between the e35S promoter and tml 3'UTR in SalI and NotI digested pMON73273. The resulting plasmid is named pMON65435.

pMON65436 (p35S::BnLDOX (sense)::tml)

The EST clone LIB4169-001-Q1-K1-E8 (SEQ ID NO: 144) is derived from the aleurone layer and seed coat of 25 to 28 days after pollination developing seed of *B. napus* (cv. Quantum). This clone, designated BnLDOX, which encodes a putative orthologue of AtLDOX, is removed by digestion with SalI and NotI. The resulting approximately 1.3 kb fragment is ligated between the e35S promoter and tml 3' UTR in SalI and NotI digested pMON73273. The resulting plasmid is named pMON65436.

pMON65437 (p35S::GmLDOX (sense)::tml)

The EST library SOYMON035 is constructed from seed coat tissue from soybean (*Glycine max*). Synthesis of cDNA is initiated using a NotI-oligo (dT) primer (SEQ ID NO: 146). Double-stranded cDNA is blunted, ligated to EcoRI adaptors, digested with NotI, size-selected, and cloned into the NotI and EcoRI sites of the pINCY vector (Incyte Genomics, Palo Alto, CA). The EST clone 701202688H1, designated GmLDOX, which encodes an orthologue of AtLDOX, is removed by digestion with EcoRI and NotI. Prior to NotI digestion, the EcoRI site is blunt ended using Pfu polymerase according to the manufacturer's instructions (Stratagene, La Jolla, CA). The resulting approximately 1.2 kb fragment is ligated between the e35S promoter and tml 3'UTR in SalI and NotI digested pMON73273. The resulting plasmid is named pMON65437.

pMON65441 (pBAN::AtLDOX (sense)::nap)

The AtLDOX gene is modified by the addition of restriction sites to facilitate cloning using primers AtLDOXNco (SEQ ID NO: 169) and AtLDOXNotI (SEQ ID NO: 170). The reaction conditions for the PCR reaction follow a protocol recommended by the enzyme manufacturer (PE Applied Biosystems, Foster City, CA). AtLDOX is amplified using LIB3177-088-P1-K1-D7 DNA as the template, 30 nanomoles each of the primers AtLDOXNco and AtLDOXNotI, 10 micromoles each of dATP, dCTP, dGTP and TTP, 2.5 units of AmpliTaq Gold (PE Applied Biosystems) in 1X Opti-Prime™ Buffer 3 (Stratagene, La Jolla, CA). After

an initial incubation at 95°C for 10 minutes, 28 cycles of PCR are performed with 94°C for 15 seconds, 62°C for 10 seconds, 52°C for 10 seconds, and 72°C for 3 minutes followed by 1 cycle of 72°C for 7 minutes. The product of the PCR reaction is purified according to standard methodology well known in the art and cloned into pCR2.1 Topo (Invitrogen) according to manufacturer's instructions. The resulting plasmid is named pMON65439. The sequence of this clone is determined using standard sequencing methodologies as set forth by PE Applied Biosystems, (Foster City, CA).

A 1070-bp fragment, containing AtLDOX, is removed by digestion with NcoI and NotI from pMON65439. The fragment is ligated between the BANYULS promoter and napin 3'UTR in NcoI and NotI digested pMON70905. The resulting plasmid is named pMON65441. The nucleic acid sequence is determined using standard sequencing methodology and confirms the integrity of the cloning junctions.

pMON69840 (pWER::AtLDOX (sense)::tml)

A 1121-bp fragment, containing AtLDOX, is removed by digestion with NcoI and BamHI from pMON65439. The BamHI overhang is blunt ended using Pfu polymerase according to the manufacturer's instructions (Stratagene, La Jolla, CA). The fragment is ligated, in place of BnTTG1, between the WEREWOLF promoter and tml 3'UTR in NcoI and PstI digested pMON65442. The PstI overhang is blunt ended using Pfu polymerase according to the manufacturer's instructions (Stratagene, La Jolla, CA). The resulting plasmid is named pMON69840. The nucleic acid sequence is determined using standard sequencing methodology and confirms the integrity of the cloning junctions.

pMON69841 (pNapin::AtLDOX (sense)::napin)

The vector pMON73274 contains a nopaline T-DNA right border sequence and octopine T-DNA left border sequence. Between the two T-DNA border sequences are the FMV.35S promoter driving expression of an CP4 EPSP synthase coding sequence consisting of the first exon of the *Arabidopsis* EPSP synthase gene (containing a chloroplast targeting sequence) linked to a synthetic EPSP synthase coding region and the 3' UTR from the pea rbcS E9 gene; the napin promoter and 3' UTR and recognition signals for cre recombinase.

A 1342-bp fragment, containing AtLDOX, is removed by digestion with SalI and NotI from pMON65435. The resulting fragment is ligated between the napin promoter and napin 3'UTR in SalI and NotI digested pMON73274. The resulting plasmid is named pMON69841.

The nucleic acid sequence is determined using standard sequencing methodology and confirms the integrity of the cloning junctions.

The resulting plasmids are introduced into *Agrobacterium tumefaciens* strain ABI using methodologies well known in the art. *Arabidopsis* plants homozygous for the *oil4* mutation are transformed and grown as described in Example 3. The seed phenotype of the transgenic *Arabidopsis* plants is compared to both mutant and wild type seed.

Twelve independently derived *oil4* plants transformed with pMON65435, pMON65436, pMON65437 or pMON65441 were grown in the same flat as six *oil4* mutants and 6 wild type Ler control plants. Seed was harvested from each individual plant and analyzed for color and seed oil and protein levels using NIR methodologies as described in Example 4. Differences were determined by a means comparison using the Student's t-test (JMP Software, SAS Institute Inc.) with $p<.05$.

Seed from plants transformed with pMON65435 did not have significantly different oil or protein levels than seeds from the untransformed *oil4* mutant nor was the seed color of transformants significantly different from the *oil4* mutant seed. Appropriate expression of AtLDOX in the transformed seed has not been confirmed.

Seed from plants transformed with pMON65436, pMON65437 and pMON65441 had significantly lower oil than seed from the untransformed *oil4* mutant. Seed from plants transformed with pMON65436 had significantly higher oil than seed from the untransformed *oil4* mutant. Although the color of seed in plants transformed with pMON65436 was indistinguishable from the mutant, seed from plants transformed with pMON65437 and pMON65441 was intermediate in color between seed from the untransformed *oil4* mutant and wild type Ler plants.

The observation of transgenic seed that phenotypically resembles wild type seed indicates complementation of the mutant by the transformed gene, that LDOX is involved in the determination of oil and protein levels, and that GmLDOX and BnLDOX encode functional homologues of the *Arabidopsis thaliana* LDOX gene.

Example 6

This example provides the construction of plant transformation vectors for the suppression of LDOX expression in canola and soybean.

LDOX suppression vectors are prepared using a vector design based upon the intron-spliced hairpin RNAs (ihpRNA) described in Smith *et al.* (*Nature* 407:319-320 (2000)). The suppression vector contains a nopaline T-DNA right border sequence and an octopine T-DNA left border sequence. Between the two T-DNA border sequences are the e35S-CMV promoter; a polylinker to facilitate cloning; the first intron from the *Arabidopsis* FAD2 gene (gi: 22655457) with flanking splice signal sequences; a second polylinker to facilitate cloning the 3' UTR of the tumor large locus (tml) from *Agrobacterium tumefaciens*; the 35S promoter from the Figwort Mosaic Virus, (35S-FMV), driving expression of an CP4 EPSP synthase coding sequence consisting of the first exon of the *Arabidopsis* EPSP synthase gene (containing a chloroplast targeting sequence) linked to a synthetic EPSP synthase coding region, and the 3' UTR from the pea rbcS E9 gene; and recognition sites for cre recombinase.

pMON69843 (e35S::BnLDOX (stabilized antisense)::tml)

The vector pMON65449 contains a nopaline T-DNA right border sequence and octopine T-DNA left border sequence, in between which are the enhanced CaMV35S promoter (P-CaMV.35S-enh-1), first intron of the *Arabidopsis* FAD2 gene, the tml 3' UTR, the P-FMV.35S-1 promoter and L-Ph.DnaK-1 leader sequence driving the expression of the CP4 EPSP synthase gene (CR-AGRtu.aroA-CP4.nno-1) containing a CTP linked to the 3' UTR from the pea rbsc E9 gene and the recognition sites for cre recombinase.

BnLDOX is removed from pMON65346 by digestion with SalI and NotI. The resulting approximately 1.4 kb fragment is ligated between the AtFAD2 intron and tml 3' UTR in SalI and PspOMI digested pMON65449. The resulting plasmid is named CLHEWI03.0045. The nucleic acid sequence was determined using standard sequencing methodology and confirmed the integrity of the cloning junctions. BnLDOX is removed from pMON65346 by digestion with SalI and NotI. The resulting approximately 1.4 kb fragment is ligated between the enhanced CaMV35S promoter and AtFAD2 intron in XhoI and NotI digested CLHEWI03.0045. The resulting plasmid is named pMON69843. The nucleic acid sequence was determined using standard sequencing methodology and confirmed the integrity of the cloning junctions

pMON69842 (Napin::BnLDOX (ihp RNA)::tml)

The vector pMON82351 contains a nopaline T-DNA right border sequence and an octopine T-DNA left border sequence, in between which are the napin promoter, first intron of the *Arabidopsis* FAD2 gene, the tml 3' UTR, the P-FMV.35S-1 promoter and L-Ph.DnaK-1 leader sequence driving the expression of the CP4 EPSP synthase gene

(CR-AGRtu.aroA-CP4.nno-1) containing a CTP linked to the 3' UTR from the pea rbsc E9 gene and the recognition sites for cre recombinase.

BnLDOX is removed from pMON65346 by digestion with SalI and NotI. The resulting approximately 1.4 kb fragment is ligated between the AtFAD2 intron and tml 3' UTR in SalI and PspOMI digested pMON82351. The resulting plasmid is named CLHEWI03.0046. The nucleic acid sequence was determined using standard sequencing methodology and confirmed the integrity of the cloning junctions. BnLDOX is removed from pMON65346 by digestion with SalI and NotI. The resulting approximately 1.4 kb fragment is ligated between the napin promoter and AtFAD2 intron in XhoI and NotI digested CLHEWI03.0046. The resulting plasmid is named pMON69842. The nucleic acid sequence was determined using standard sequencing methodology and confirmed the integrity of the cloning junctions

pMON65446 (BAN::BnLDOX (ihp RNA)::tml)

The vector pMON65443 contains a nopaline T-DNA right border sequence and an octopine T-DNA left border sequence, in between which are the Banyuls promoter, first intron of the *Arabidopsis* FAD2 gene, the tml 3' UTR, the P-FMV.35S-1 promoter and L-Ph.DnaK-1 leader sequence driving the expression of the CP4 EPSP synthase gene (CR-AGRtu.aroA-CP4.nno-1) containing a CTP linked to the 3' UTR from the pea rbsc E9 gene and the recognition sites for cre recombinase.

BnLDOX is removed from EST clone LIB4169-001-Q1-K1-E8 (SEQ ID NO: 144) by digestion with SalI and NotI. The resulting approximately 1.4 kb fragment is ligated between the AtFAD2 intron and tml 3' UTR in SalI and PspOMI digested pMON65443. The resulting plasmid is named DMRUEZ03.0053. The nucleic acid sequence was determined using standard sequencing methodology and confirmed the integrity of the cloning junctions. BnLDOX is removed from EST clone LIB4169-001-Q1-K1-E8 (SEQ ID NO: 144) by digestion with SalI and NotI. The resulting approximately 1.4 kb fragment is ligated between the banyuls promoter and AtFAD2 intron in XhoI and NotI digested DMRUEZ03.0053. The resulting plasmid is named pMON65446 (Figure 4). The nucleic acid sequence was determined using standard sequencing methodology and confirmed the integrity of the cloning junctions

pMON65447 (WER::BnLDOX (ihp RNA)::tml)

The vector pMON65444 contains a nopaline T-DNA right border sequence and octopine T-DNA left border sequence, in between which are the WEREWOLF promoter, first intron of the Arabidopsis FAD2 gene, the tml 3' UTR, the P-FMV.35S-1 promoter and

L-Ph.DnaK-1 leader sequence driving the expression of the CP4 EPSP synthase gene (CR-AGRtu.aroA-CP4.nno-1) containing a CTP linked to the 3' UTR from the pea rbsc E9 gene and the recognition sites for cre recombinase.

BnLDOX is removed from EST clone LIB4169-001-Q1-K1-E8 (SEQ ID NO: 144) by digestion with SalI and NotI. The resulting approximately 1.4 kb fragment is ligated between the AtFAD2 intron and tml 3' UTR in SalI and PspOMI digested pMON65444. The resulting plasmid is named DMRUEZ03.0054. The nucleic acid sequence was determined using standard sequencing methodology and confirmed the integrity of the cloning junctions. BnLDOX is removed from EST clone LIB4169-001-Q1-K1-E8 (SEQ ID NO: 144) by digestion with SalI and NotI. The resulting approximately 1.4 kb fragment is ligated between the banyuls promoter and AtFAD2 intron in XhoI and NotI digested DMRUEZ03.0054. The resulting plasmid is named pMON65447. The nucleic acid sequence was determined using standard sequencing methodology and confirmed the integrity of the cloning junctions

pMON69847 (e35S::GmLDOX (ihp RNA)::tml)

The vector pMON82359 contains a nopaline T-DNA right border sequence and octopine T-DNA left border sequence, in between which are the promoter, leader sequence and intron from the *Arabidopsis* actin-7 gene, driving the expression of the *Agrobacterium tumefaciens* CP4 EPSP synthase gene, containing a CTP, and the 3' UTR from the pea rbsc E9 gene. A 3810-bp fragment containing the enhanced CaMV 35S promoter, GmLDOX and tml 3' UTR is removed by digestion with PacI and AscI from pMON65437. The resulting fragment is ligated between the 3' UTR from the pea rbsc E9 gene and the octopine T-DNA left border sequence in PacI and AscI digested pMON82359. The resulting plasmid is named CLHEWI03.0055. The nucleic acid sequence was determined using standard sequencing methodology and confirmed the integrity of the cloning junctions. An 1156-bp fragment containing part of the GmLDOX gene is removed by digestion with StuI and NotI from pMON65437. The NotI overhang was blunt ended using Pfu polymerase according to the manufacturer's instructions (Stratagene, La Jolla, CA). The resulting fragment was ligated between the AtFAD2 intron and tml 3' UTR in SmaI digested pMON65449. The resulting plasmid is named CLHEWI03.0059. The nucleic acid sequence was determined using standard sequencing methodology and confirmed the integrity of the cloning junctions. A 2410 bp fragment containing the AtFAD2 intron and GmLDOX is removed by SacI and NotI digestion of CLHEWI03.0059. The SacI overhang is blunt ended using Pfu polymerase according to the manufacturer's instructions (Stratagene, La Jolla, CA). The resulting fragment is ligated

between GmLDOX and the tml3' UTR in SwaI and NotI digested CLHEWI03.0055. The resulting plasmid is named pMON69847 (Figure 5). The nucleic acid sequence was determined using standard sequencing methodology and confirmed the integrity of the cloning junctions.

Additional suppression vectors are constructed by substituting tissue specific promoters such as pBANYULs or pWEREWOLF for the e35S promoter. These constructs allow for tissue specific suppression of the LDOX orthologues. Table 2 summarizes the vectors constructed. Canola plants are transformed with the vectors as described in Example 3.

Table 2. Vectors constructed (ihpRNA = intron-hairpin RNA).

Construct ID	Promoter	Gene of Interest	Vector Type	Intron	3' UTR
pMON69843	e35S	BnLDOX	ihpRNA	AtFAD2	tml
pMON69841	Napin	BnLDOX	ihpRNA	AtFAD2	tml
pMON65446	Banyuls	BnLDOX	ihpRNA	AtFAD2	tml
pMON65447	WEREWOLF	BnLDOX	ihpRNA	AtFAD2	tml
pMON69847	e35S	GmLDOX	ihpRNA	AtFAD2	tml
pMONXXXXX	Banyuls	GmLDOX	ihpRNA	AtFAD2	tml
pMONXXXXX	WEREWOLF	GmLDOX	ihpRNA	AtFAD2	tml

Example 7

This example describes the identification of orthologues to the *Arabidopsis* CHS, LDOX and TTG1 genes.

The sequence for the *Arabidopsis* TTG1 protein (gi: 5123716, SEQ ID NO: 1) was used to BLAST search against 6 frame translations of proprietary soybean and corn cDNA libraries using the tBLASTn protocol. The cDNA sequence of the top 2-4 hits were then used to BLAST search against a proprietary *Arabidopsis thaliana* sequence database. Both the soybean cDNA (700727328_FLI, SEQ ID NO: 2) and the corn cDNA (ZEAMA-22JAN02-CLUSTER6987_1, SEQ ID NO: 3) had TTG1 as their top hit in *Arabidopsis*, thus confirming the soybean and corn sequences as orthologs of TTG1.

The sequence for the *Arabidopsis* CHS (gi: 23308391, SEQ ID NO: 168) was used to BLAST search against six frame translations of proprietary soybean and corn cDNA libraries. Five corn and three soy ESTs had an expectation value of 0.0. These eight ESTs (SEQ ID NOs: 4-11) were used to BLAST search against a proprietary *Arabidopsis* sequence database. The top hit for each of the sequences was the *Arabidopsis* chalcone synthase gene, thus confirming the soybean and corn sequences as orthologs for *Arabidopsis* CHS.

The sequence of corn cDNA (ZEAMA-22JAN02-CLUSTER6987_1, SEQ ID NO: 3) was used as a BLAST query against proprietary and public *Zea mays* sequence databases. Two sequences were identified as having high homology in proprietary databases (SEQ ID NOs: 12 and 13). There were no sequences in public databases that were identified as having high homology.

The sequence of soybean cDNA (700727328_FL1, SEQ ID NO: 2) was used as a BLAST query against proprietary and public *Glycine max* sequence databases. Four sequences were identified as having high homology in proprietary databases (SEQ ID NOs: 14-17). Eleven sequences were identified as having high homology in public databases (SEQ ID NOs: 18-28). Table 3 shows the NCBI annotations for the public orthologs.

Table 3. Annotations for public orthologs.

SEQ ID NO	NCBI Description	Genbank sequence ID
18	similar to TR:Q9XGN1 Q9XGN1 TTG1 PROTEIN.	19936077
19	similar to TR:O24514 O24514 AN11.	4303493
20	similar to TR:Q9XGN1 Q9XGN1 TTG1 PROTEIN.	18848414
21	similar to TR:O24514 O24514 AN11.	15000120
22	similar to TR:Q9XGN1 Q9XGN1 TTG1 PROTEIN.	14991581
23	similar to TR:O24514 O24514 AN11.	18040199
24	similar to TR:Q9XGN1 Q9XGN1 TTG1 PROTEIN.	10843406
25	similar to TR:Q9XGN1 Q9XGN1 TTG1 PROTEIN.	14011464
26	similar to TR:O24514 O24514 AN11.	4307857
27	similar to TR:O24514 O24514 AN11.	15203321
28	similar to TR:O24514 O24514 AN11.	10844218

The sequence of *Brassica napus* cDNA (LIB3169-025-P1-K2-B5, SEQ ID NO: 30) was used as a BLAST query against proprietary and public *Brassica napus* databases. One sequence was identified as having high homology in proprietary databases (SEQ ID NO: 29). There were no sequences identified in the public databases searched as having high homology. The predicted amino acid sequence for the *Arabidopsis* LDOX protein (SEQ ID NO: 124) was used as a query in a BLAST search, using the tBLASTn protocol, against both public (non-redundant) and proprietary databases. Fourteen sequences were identified as having high homology in the proprietary databases (SEQ ID NOs: 127-140). Table 4 shows the results of the searches.

Table 4. *Brassica napus* sequences.

SEQ ID NO:	DESCRIPTION	ANNOTATION
127	ARATH-01JUL02-CLUSTER5762_1	leucoanthocyanidin dioxygenase (EC 1.14.11.-)
128	HORVU-20MAR02-CLUSTER23174_1	AB073919) anthocyanidin synthase [Ipomoea nil]
129	BRANA-18JUN02-CLUSTER39879_1	leucoanthocyanidin dioxygenase (EC 1.14.11.-)
130	LIB80-012-Q1-E1-C7	leucoanthocyanidin dioxygenase (EC 1.14.11.-)
131	ZEAMA-06JUN02-CLUSTER2276_1	leucoanthocyanidin dioxygenase (LDOX)
132	GOSHI-09MAY01-CLUSTER11019_1	
133	LIB3135-036-Q1-K1-E8	leucoanthocyanidin dioxygenase (EC 1.-.-.-) (LDOX)
134	LIB3146-009-Q1-K1-A4	leucoanthocyanidin dioxygenase (LDOX)
135	ALLPO-19APR01-CLUSTER7272_2	(AB011796) flavonol synthase [*Citrus unshiu*]
136	ALLPO-19APR01-CLUSTER7272_1	(AB011796) flavonol synthase [*Citrus unshiu*]
137	ORYSA-27FEB02-CLUSTER915_1	leucoanthocyanidin dioxygenase (EC 1.14.11.-)
138	ORYSA-27FEB02-CLUSTER84577_1	(AB026295) Similar to leucoanthocyanidin dioxygenase
139	SORBI-27FEB02-CLUSTER59710_1	(AC069300) putative dioxygenase [*Oryza sativa*]
140	GLYMA-06JUN02-CLUSTER7806_1	leucoanthocyanidin dioxygenase (EC 1.14.11.-)

Example 8

This example sets forth the identification and sequence analysis of PAL orthologs from corn, soybean and *Brassica.*

Sequences for the three forms of phenylalanine ammonia lyase from *Arabidopsis* were retrieved from the NCBI database by entering a search using the keywords PAL1, PAL2 and PAL3. The nucleic acid sequences for these three enzymes are given in SEQ ID NO: 151-153, respectively. These three sequences were then used to BLAST search against proprietary databases for *Arabidopsis*, soybean, and *Brassica*. Only one hit was identified when using the BLASTN 2.0.12 protocol. This sequence, LIB4315-024-R1-N1-B9, from the corn database, is shown in SEQ ID NO: 154. Six frame translations were then BLAST searched against six frame translations of the corn, soybean and *Brassica* databases using the TBLASTX 2.0.12 protocol. The top five hits from the corn and rice databases are shown in Table 5.

Table 5. PAL orthologs.

Query	Database	SEQ ID NO:	Identification
AtPAL1	CornUnigeneDNA	155	ZEAMA-06JUN02-CLUSTER1962_14
AtPAL1	CornUnigeneDNA	156	ZEAMA-06JUN02-CLUSTER1962_42
AtPAL1	CornUnigeneDNA	157	ZEAMA-06JUN02-CLUSTER1962_26
AtPAL1	CornUnigeneDNA	158	ZEAMA-06JUN02-CLUSTER1962_30
AtPAL1	CornUnigeneDNA	159	ZEAMA-06JUN02-CLUSTER1962_27
AtPAL2	CornUnigeneDNA	155	ZEAMA-06JUN02-CLUSTER1962_14
AtPAL2	CornUnigeneDNA	156	ZEAMA-06JUN02-CLUSTER1962_42
AtPAL2	CornUnigeneDNA	159	ZEAMA-06JUN02-CLUSTER1962_27
AtPAL2	CornUnigeneDNA	157	ZEAMA-06JUN02-CLUSTER1962_26
AtPAL2	CornUnigeneDNA	158	ZEAMA-06JUN02-CLUSTER1962_30
AtPAL3	CornUnigeneDNA	158	ZEAMA-06JUN02-CLUSTER1962_30
AtPAL3	CornUnigeneDNA	157	ZEAMA-06JUN02-CLUSTER1962_26
AtPAL3	CornUnigeneDNA	159	ZEAMA-06JUN02-CLUSTER1962_27
AtPAL3	CornUnigeneDNA	160	ZEAMA-06JUN02-CLUSTER1962_36
AtPAL3	CornUnigeneDNA	155	ZEAMA-06JUN02-CLUSTER1962_14
AtPAL1	RiceUnigeneDNA	161	ORYSA-27FEB02-CLUSTER243_23
AtPAL1	RiceUnigeneDNA	162	ORYSA-27FEB02-CLUSTER243_24
AtPAL1	RiceUnigeneDNA	163	ORYSA-27FEB02-CLUSTER243_13
AtPAL1	RiceUnigeneDNA	164	ORYSA-27FEB02-CLUSTER243_9
AtPAL1	RiceUnigeneDNA	165	ORYSA-27FEB02-CLUSTER243_19
AtPAL2	RiceUnigeneDNA	161	ORYSA-27FEB02-CLUSTER243_23
AtPAL2	RiceUnigeneDNA	162	ORYSA-27FEB02-CLUSTER243_24
AtPAL2	RiceUnigeneDNA	163	ORYSA-27FEB02-CLUSTER243_13
AtPAL2	RiceUnigeneDNA	164	ORYSA-27FEB02-CLUSTER243_9
AtPAL2	RiceUnigeneDNA	165	ORYSA-27FEB02-CLUSTER243_19
AtPAL3	RiceUnigeneDNA	162	ORYSA-27FEB02-CLUSTER243_24
AtPAL3	RiceUnigeneDNA	161	ORYSA-27FEB02-CLUSTER243_23
AtPAL3	RiceUnigeneDNA	163	ORYSA-27FEB02-CLUSTER243_13
AtPAL3	RiceUnigeneDNA	164	ORYSA-27FEB02-CLUSTER243_9
AtPAL3	RiceUnigeneDNA	165	ORYSA-27FEB02-CLUSTER243_19

Example 9

This example describes the construction of plant transformation vectors containing PAL orthologs.

PAL suppression vectors are prepared using a vector design based upon the intron-spliced hairpin RNAs described in Smith *et al.* (*Nature* 407:319-320 (2000)). The suppression vector contains a nopaline T-DNA right border sequence and an octopine T-DNA left border sequence. Between the two T-DNA border sequences are the e35S-CMV promoter;

a polylinker to facilitate cloning; the first intron from the *Arabidopsis* FAD2 gene (gi: 22655457) with flanking splice signal sequences; a second polylinker to facilitate cloning the 3' UTR of the tumor large locus (tml) from *Agrobacterium tumefaciens*; the 35S promoter from the Figwort Mosaic Virus (35S-FMV) driving expression of an CP4 EPSP synthase coding sequence consisting of the first exon of the *Arabidopsis* EPSP synthase gene (containing a chloroplast targeting sequence) linked to a synthetic EPSP synthase coding region, and the 3' UTR from the pea rbcS E9 gene; and recognition sites for cre recombinase.

Gene segments containing at least 100 base pairs of one or more *Brassica* PAL orthologs are cloned in the opposite (antisense) orientation on either side of the AtFAD2 intron in the vector described above. Additional suppression vectors are constructed by substituting tissue-specific promoters such as pBANYULs or pWEREWOLF for the e35S promoter. These constructs allow for tissue-specific suppression of the PAL orthologs. Similar suppression vectors containing sequences suitable for soybean transformation and soybean PAL orthologs are prepared for suppression of PAL orthologs in soybean.

pMON82708 (pBAN::PAL1-PAL2-PAL3 (ihp RNA)::tml)

Gene fragments from *Brassica napus* PAL genes are isolated by PCR using primers: PAL_A_FWD_19617 (SEQ ID NO: 171); PAL_A_REV_19618 (SEQ ID NO: 172); PAL_B_FWD_19619 (SEQ ID NO: 173); PAL_B_REV_19620 (SEQ ID NO: 174); PAL_C_FWD_19621 (SEQ ID NO: 175); and PAL_C_Rev_19622 (SEQ ID NO: 176).

The reaction conditions for the PCR reaction follow a protocol recommended by the enzyme manufacturer (PE Applied Biosystems, Foster City, CA). BnPAL1 is amplified using LIB3032-008-Q1-E1-H7 as a template with 30 nanomoles each of the primers PAL_B_FWD_19619 (SEQ ID NO: 173) and PAL_B_REV_19620 (SEQ ID NO: 174), using 10 micromoles each of dATP, dCTP, dGTP and TTP, 2.5 units of AmpliTaq Gold (PE Applied Biosystems) in 1X Opti-Prime™ Buffer 3 (Stratagene). After an initial incubation at 95°C for 10 minutes, 28 cycles of PCR are performed with 94°C for 15 seconds, 62°C for 10 seconds, 52°C for 10 seconds, and 72°C for 3 minutes followed by 1 cycle of 72°C for 7 minutes. The product of the PCR reaction is purified according to standard methodology well known in the art and cloned into pCR2.1 Topo (Invitrogen) according to manufacturer's instructions. The resulting plasmid is designated pBnPAL1. The sequence of this clone was determined using standard sequencing methodologies as set forth by PE Applied Biosystems, (Foster City, CA).

BnPAL2 is amplified using LIB4169-009-Q1-K1-B3 as a template with 30 nanomoles each of the primers PAL_C_FWD_19621 (SEQ ID NO: 175) and PAL_C_Rev_19622 (SEQ ID NO: 176), 10 micromoles each of dATP, dCTP, dGTP and TTP, 2.5 units of AmpliTaq Gold (PE Applied Biosystems) in 1X Opti-Prime™ Buffer 3 (Stratagene). After an initial incubation at 95°C for 10 minutes, 28 cycles of PCR are performed with 94°C for 15 seconds, 62°C for 10 seconds, 52°C for 10 seconds, and 72°C for 3 minutes followed by 1 cycle of 72°C for 7 minutes. The product of the PCR reaction is purified according to standard methodology well known in the art and cloned into pCR2.1 Topo (Invitrogen) according to manufacturer's instructions. The resulting plasmid is designated pBnPAL2. The sequence of this clone was determined using standard sequencing methodologies as set forth by PE Applied Biosystems, (Foster City, CA).

BnPAL3 is amplified using LIB4153-016-R1-K1-D2 as a template with 30 nanomoles each of the primers PAL_A_FWD_19617 (SEQ ID NO: 171) and PAL_A_REV_19618 (SEQ ID NO: 172), 10 micromoles each of dATP, dCTP, dGTP and TTP, 2.5 units of AmpliTaq Gold (PE Applied Biosystems) in 1X Opti-Prime™ Buffer 3 (Stratagene). After an initial incubation at 95°C for 10 minutes, 28 cycles of PCR were performed with 94°C for 15 seconds, 62°C for 10 seconds, 52°C for 10 seconds, and 72°C for 3 minutes followed by 1 cycle of 72°C for 7 minutes. The product of the PCR reaction is purified according to standard methodology well known in the art and cloned into pCR2.1 Topo (Invitrogen) according to manufacturer's instructions. The resulting plasmid is designated pBnPAL3. The sequence of this clone was determined using standard sequencing methodologies as set forth by PE Applied Biosystems (Foster City, CA).

BnPAL2 is removed from pBnPAL2 by digestion with AatII and KpnI. The resulting fragment is ligated next to BnPAL1 in AatII and KpnI digested pBnPAL1 to create pBNPal12. BnPAL3 is removed from pBnPAL3 by digestion with EagI. The resulting fragment is ligated next to BnPAL1 in NotI digested pBnPAL12 to create PDROBER03.0020.

The nucleic acid sequence was determined using standard sequencing methodology and confirmed the integrity of the cloning junctions.

A 1053 bp fragment containing BnPAL1, BnPAL2 and BnPAL3 is removed from PDROBER03.0020 by digestion with BamHI and XhoI. The BamHI overhang is blunt ended using standard methodology. The resulting fragment is ligated in between the AtFAD2 intron and tml 3' UTR in SalI and SmaI digested pMON65443. The resulting plasmid is named

PDROBER03.0023. The nucleic acid sequence was determined using standard sequencing methodology and confirmed the integrity of the cloning junctions. A 1053 bp fragment containing BnPAL1, BnPAL2 and BnPAL3 is removed from PDROBER03.0020 by digestion with BamHI and XhoI. The BamHI overhang is blunt ended using standard methodology. The resulting fragment is ligated in between the banyuls promoter and AtFAD2 intron in XhoI and NotI digested PDROBER03.0023. The NotI overhang is blunt ended using standard methodology. The resulting plasmid is named pMON82708 (Figure 6). The nucleic acid sequence was determined using standard sequencing methodology and confirmed the integrity of the cloning junctions.

<u>pMON82709 e35S::PAL1-PAL2-PAL3 (ihp RNA)::tml)</u>

A 1053 bp fragment containing BnPAL1, BnPAL2 and BnPAL3 is removed from PDROBER03.0020 by digestion with BamHI and XhoI. The BamHI overhang is blunt ended using standard methodology. The resulting fragment is ligated in between the AtFAD2 intron and tml 3' UTR in SalI and SmaI digested pMON65449. The resulting plasmid is named PDROBER03.0025. The nucleic acid sequence is determined using standard sequencing methodology and confirmed the integrity of the cloning junctions. A 1053 bp fragment containing BnPAL1, BnPAL2 and BnPAL3 is removed from PDROBER03.0020 by digestion with BamHI and XhoI. The BamHI overhang is blunt ended using standard methodology. The resulting fragment is ligated in between the enhanced 35S promoter and AtFAD2 intron in XhoI and NotI digested PDROBER03.0025. The NotI overhang is blunt ended using standard methodology. The resulting plasmid is named pMON82709. The nucleic acid sequence was determined using standard sequencing methodology and confirmed the integrity of the cloning junctions. Table 6 summarizes several of the vectors described above. Canola plants are transformed with the vectors as described in Example 3.

Table 6. Vector constructs (ihpRNA = intron-hairpin RNA).

Construct ID	Promoter	Gene of Interest	Vector Type	Intron	3' UTR
pMONXXXXX	e35S	BnPAL1	ihpRNA	AtFAD2	tml
pMONXXXXX	Banyuls	BnPAL1	ihpRNA	AtFAD2	tml
pMONXXXXX	WEREWOLF	BnPAL1	ihpRNA	AtFAD2	tml
pMONXXXXX	e35S	GmPAL1	ihpRNA	AtFAD2	tml
pMONXXXXX	Banyuls	GmPAL1	ihpRNA	AtFAD2	tml
pMONXXXXX	WEREWOLF	GmPAL1	ihpRNA	AtFAD2	tml
pMONXXXXX	e35S	BnPAL2	ihpRNA	AtFAD2	tml
pMONXXXXX	Banyuls	BnPAL2	ihpRNA	AtFAD2	tml

Construct ID	Promoter	Gene of Interest	Vector Type	Intron	3' UTR
pMONXXXXX	WEREWOLF	BnPAL2	ihpRNA	AtFAD2	tml
pMONXXXXX	e35S	GmPAL2	ihpRNA	AtFAD2	tml
pMONXXXXX	Banyuls	GmPAL2	ihpRNA	AtFAD2	tml
pMONXXXXX	WEREWOLF	GmPAL2	ihpRNA	AtFAD2	tml
pMONXXXXX	e35S	BnPAL3	ihpRNA	AtFAD2	tml
pMONXXXXX	Banyuls	BnPAL3	ihpRNA	AtFAD2	tml
pMONXXXXX	WEREWOLF	BnPAL3	ihpRNA	AtFAD2	tml
pMONXXXXX	e35S	GmPAL3	ihpRNA	AtFAD2	tml
pMONXXXXX	Banyuls	GmPAL3	ihpRNA	AtFAD2	tml
pMONXXXXX	WEREWOLF	GmPAL3	ihpRNA	AtFAD2	tml
pMON82709	e35S	BnPAL1-PAL2-PAL3	ihpRNA	AtFAD2	tml
pMON82708	Banyuls	BnPAL1-PAL2-PAL3	ihpRNA	AtFAD2	tml
pMONXXXXX	WEREWOLF	BnPAL1-PAL2-PAL3	ihpRNA	AtFAD2	tml
pMONXXXXX	e35S	GmPAL1-PAL2-PAL3	ihpRNA	AtFAD2	tml
pMONXXXXX	Banyuls	GmPAL1-PAL2-PAL3	ihpRNA	AtFAD2	tml
pMONXXXXX	WEREWOLF	GmPAL1-PAL2-PAL3	ihpRNA	AtFAD2	tml

Example 10

This example describes complementation of *Arabidopsis* mutants transformed with pMON65442 or pMON65440 as described in Example 3.

One proof of successful identification and cloning involves complementation of the mutant phenotype with the wild type allele (Lui *et al., Proc. Natl. Acad. Sci. (U.S.A.,)* 96:6535–6540 (1999)). Furthermore, Dong *et al.* established that complementation of *Arabidopsis* transparent testa mutants should be a useful system for establishing the function of genes with homology to flavonoid biosynthetic enzymes (Dong *et al., Plant Physiology,* 127:46-57 (2001)), and Nesi *et al.* used genetic complementation to confirm the identification of the gene encoded by TT16, a transcription factor that regulates genes in the flavonoid biosynthetic pathway (Nesi *et al., Plant Cell,* 14:2463-2479 (2002)).

By expressing a functional TTG1 gene in a tissue-specific manner, cells or organs where expression of TTG1 influences seed composition can be identified. Twelve independently derived ttg1-1 plants transformed with pMON65442 or pMON65440 were grown in the same flat as six ttg1-1 mutants and 6 wild type Ler control plants. Seed was harvested from each individual plant and analyzed for color and seed oil and protein levels using NIR methodologies as described in Example 4. Seed from plants transformed with pMON65442 had significantly lower oil levels than seeds from the untransformed ttg1-1 mutant (38.7% vs. 44.2%). Wild type

Ler plants produced seed with an average oil level of 41.4%. pMON65442 did not complement the high protein or seed color phenotypes of the ttg1-1 mutant. This shows that expression of TTG1 in the root atrichoblast is involved in determining the oil and protein levels in seeds.

Seedlings were also scored for the presence of trichomes on true leaves. From 80 transformants tested, all had trichomes. The transformants have brown seeds and significantly lower protein levels than untransformed ttg-1 plants, and equivalent levels to the wild type, as determined by a means comparison using the Student's t-test (JMP Software, SAS Institute Inc.). For wild type, the average protein level was 21.5%. For untransformed ttg-1 plants, the average protein level was 25%. For transformants, the average protein level was 23%. These results show that the BnTTG1 constructs complemented the mutant ttg-1 plants and restored oil and protein to wild type levels. This shows that TTG1 is involved in determining the levels of oil and protein in seeds and that BnTTG1 is a functional homologue of the *Arabidopsis* TTG1 gene.

For pMON65440, successful complementation could not be determined because no significant difference in oil was detected between the control wild type and mutant plants due to a high level of variation. Complementation tests for pMON65440 are being repeated.

Example 11

This example describes results in canola transformed by methods in Example 3 with constructs from Example 2.

Analysis of Canola plants transformed with pMON75051

For each event tested, seventeen gene of interest positive (transgenic) and seventeen null segregants (control) R1 plants derived from the same initial transformant were grown. Seed was harvested from each individual plant and analyzed for oil and protein levels using NIR methodologies as described in Example 4 (Table 7). Differences between transgenic and control plants for each event were determined by a means comparison using the Student's t-test (JMP Software, SAS Institute Inc.) with $p<.05$. No significant differences in oil level were found between transgenic and control seed. Transgenic seed for plants for one event, BN_G942, had a significantly higher level of protein.

Table 7. Average R2 seed oil and protein levels of plants transformed with pMON75051

EVENT	% OIL		% PROTEIN	
	Control	Transgenic	Control	Transgenic
BN_G895	39.4	40.0	25.2	24.9
BN_G910	39.4	39.6	28.2	27.8
BN_G913	39.5	39.5	26.8	26.8
BN_G918	38.8	39.6	27.8	27.7
BN_G942	39.6	39.1	27.0	27.4

<u>Analysis of Canola plants transformed with pMON75052</u>

For each of five individual events tested, seventeen gene of interest positive (transgenic) and seventeen null segregants (control) R1 plants derived from the same initial transformant were grown. Seed was harvested from each individual plant and analyzed for oil and protein levels using NIR methodologies as described in Example 4 (Table 8). Differences between transgenic and control plants for each event were determined by a means comparison using the Student's t-test (JMP Software, SAS Institute Inc.) with $p<.05$. No significant differences in oil or protein level were found between transgenic and control seed.

Table 8. Average R2 seed oil and protein levels of plants transformed with pMON75052

EVENT	% OIL		% PROTEIN	
	Control	Transgenic	Control	Transgenic
BN_G950	40.6	40.6	26.1	26.0
BN_G965	42.7	42.6	26.1	26.1
BN_G966	40.2	41.3	27.8	27.5
BN_G975	41.2	40.3	24.6	24.8
BN_G979	40.3	40.0	26.8	27.0

<u>Analysis of Canola plants transformed with pMON75702</u>

For each of six individual events tested, seventeen gene of interest positive (transgenic) and seventeen null segregants (control) R1 plants derived from the same initial transformant, were grown. Seed was harvested from each individual plant and analyzed for oil and protein levels using NIR methodologies as described in Example 4 (Table 9). Differences between transgenic and control plants for each event were determined by a means comparison using the Student's t-test (JMP Software, SAS Institute Inc.) with $p<.05$. No significant differences in oil or protein level were found between transgenic and control seed.

Table 9. Average R2 seed oil and protein levels of plants transformed with pMON75702

EVENT	Leaf Trichomes	% OIL		% PROTEIN	
		Control	Transgenic	Control	Transgenic
BN_G1439	Absent	44.9	44.5	23.1	23.1
BN_G1441	Absent	42.6	43.0	21.7	22.0
BN_G1524	Absent	43.3	43.1	21.6	21.9
BN_G1550	Absent	43.7	43.3	23.5	24.2
BN_G1602	Absent	42.5	43.3	24.2	23.6
BN_G1604	Absent	42.0	42.1	24.5	24.0

Herbicide-resistant seedlings from all six events had no detectable leaf trichomes, whereas herbicide-sensitive segregants of the same event had normal leaf trichomes. This indicates that TTG1 activity is suppressed in leaf trichome forming cells. There were no visually detectable differences in color between seed harvested from herbicide resistant and non resistant segregants.

To determine whether TTG1 was suppressed in transgenic plants, RNA isolated from developing seed and leaf was analyzed for expression levels of BnTTG1 by real time RT-PCR. Plants with a significant decrease in TTG1 transcript when compared to wild type could not be identified. Developing seed from pMON75702 transformants was screened for expression of BnLDOX, a potential target of TTG1. Plants with a significant decrease in LDOX transcript when compared to wild type could not be identified.

Analysis of Canola plants transformed with pMON70906

For each of five individual events tested, seventeen gene of interest positive (transgenic) and seventeen null segregants (control) R1 plants derived from the same initial transformant, were grown. Seed was harvested from each individual plant and analyzed for oil and protein levels using NIR methodologies as described in Example 4 (Table 10). Differences between transgenic and control plants for each event were determined by a means comparison using the Student's t-test (JMP Software, SAS Institute Inc.) with $p<.05$. No significant differences in oil level were found between transgenic and control seed. Transgenic seed for plants for one event, BN_G1303, had a significantly lower level of protein.

Table 10. Average R2 seed oil and protein levels of plants transformed with pMON70906.

EVENT	% OIL		% PROTEIN	
	Control	Transgenic	Control	Transgenic
BN_G1234	44.3	44.3	23.8	23.7
BN_G1235	42.6	42.6	23.6	23.5
BN_G1243	44.1	44.0	22.9	23.0
BN_G1264	43.6	43.9	22.9	23.0
BN_G1303	40.3	40.8	26.4	26.0

To determine whether TTG1 was suppressed in transgenic plants, RNA isolated from developing seed was analyzed for expression levels of BnTTG1 by real time RT-PCR. Plants with a significant decrease in TTG1 transcript when compared to wild type could not be identified. RNA isolated from developing whole seed, embryos and seed coats of pMON70906 transformants was screened for expression of BnLDOX, a potential target of TTG1. Plants with a significant decrease in LDOX transcript when compared to wild type could not be identified.

Extracts from transgenic and control seed, seed coat, root or vegetative tissue will be analyzed by HPLC, TLC, histochemical staining or spectrophotometric assay for phenylpropanoids using methods known to those skilled in the art. Morphological analysis to determine whether epidermal cell development has been altered, particularly in the seed coat and root, is also performed. Plants with reduced levels of phenylpropanoids are positively correlated to seed with increased levels of oil or protein.

To facilitate down regulation of the phenylpropanoid pathway, additional suppression constructs are introduced into transgenic plants. Preferred polynucleotide sequences used in the context of down regulating the phenylpropanoid pathway include those that encode a phenylpropanoid pa5thway gene, including, for example, TTG1, CHS, PAL, and LDOX. More preferably such polynucleotides include SEQ ID NOs: 2-17, 29, 30, 32-96, 121-123, 127-141, 144, 145, 147-165, and 168. A construct capable of forming a ds RNA containing coding or noncoding sequences from the aforementioned TTG1, CHS, PAL, and/or LDOX genes under control of a constitutive (such as CaMV 35S or FMV) or tissue specific (such as a root atrichoblast, seed coat, or seed) promoter is introduced into canola, soy, or corn by methods known to one skilled in the art, such as those described in Example 3.